THE WORLD BANK

Research Observer

THE WORLD BANK

Research Observer

The World Bank Research Observer is intended for anyone who has a professional interest in development. *Observer* articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the *Observer* considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the editor. Please direct all editorial correspondence to the Editor, *The World Bank Research Observer*, at the address in the copyright notice below.

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

The World Bank Research Observer is published twice a year in February and August. Single copies may be purchased for US$12.95. Subscription rates are as follows:

	Individuals	Institutions
1-year subscription	US$25	US$40
2-year subscription	US$45	US$75
3-year subscription	US$60	US$100

Orders should be sent to: World Bank Publications, Box 7247-7956, Philadelphia, Pennsylvania 19170-7956, U.S.A. Readers with mailing addresses in developing countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request. Written confirmation is required every two years to renew such subscriptions.

ISBN 0-8213-4610-5, ISSN 0257-3032

The World Bank Research Observer is indexed by the *Journal of Economic Literature*, the *Standard Periodical Directory*, ABI/INFORM, the Index to International Statistics, the Public Affairs Information Service, Inc., the Social Sciences Citation Index, and, online, by the Economic Literature Index and DIALOG. It is available in microform through University Microfilms, Inc., 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.

THE WORLD BANK

Research Observer

Volume 15 • Number 1 • February 2000

A SYMPOSIUM ON FIRM AND INDUSTRY PERFORMANCE IN THE EAST ASIAN CRISIS

THE WORLD BANK

Research Observer

Volume 15 • Number 2 • August 2000

Crisis, Adjustment, and Reform in Thailand's Industrial Firms

David Dollar • Mary Hallward-Driemeier

New data on Thailand's industrial firms shed light on the origins of the East Asian financial crisis and on the response of the manufacturing sector to the structural adjustment program supported by the international financial institutions. Before the crisis, Thai firms had declining profitability, but they nevertheless maintained high levels of investment, often in domestically oriented areas (notably the auto sector). Thai firms financed these investments with short-term borrowing from financial institutions, which in turn borrowed short term on foreign markets. That only 40 percent of firms provided audited financial statements to their banks meant that the financial sector had poor information for assessing the true riskiness of these investments. The financial structure was thus vulnerable even to small shocks.

How well did the adjustment program deal with the crisis? Thai firms had difficulty increasing their exports quickly because of investment in the wrong sectors, a decline in regional demand, and bottlenecks that included red tape and poor customs administration. Because of the poor export response, the brunt of adjustment had to come through compression of demand and of imports. In retrospect, the macroeconomic program—which assumed quick export recovery—was too tight.

Devaluation is not the policy of the Thai government.
— Japanese auto parts producer, May 1997

The anonymous entrepreneur quoted above borrowed $40 million—short term—in Singapore to finance a factory in Thailand that produced auto parts for a Japanese assembler selling in the Thai market. Taking on the debt seemed like a good idea to the Japanese producer in late 1996 and early 1997. But by October 1997, with the Thai baht devalued and the region in recession, the entrepreneur had halted production and was waiting for the ax to fall.

The World Bank Research Observer, vol. 15, no. 1 (February 2000), pp 1–22

The currency and financial crises in East Asia were touched off when Thailand devalued the baht on July 2, 1997. Within weeks Malaysia and Indonesia devalued their currencies, and the Philippine and Korean currencies also began to weaken. Stock markets across the region fell as investors pulled out their capital. Thailand quickly reached an agreement with the International Monetary Fund (IMF) for a bailout package and in return initiated the recommended reform measures, tightening monetary and fiscal policies. Although the aim was to bolster the exchange rate and to restore investor confidence, the rise in interest rates, the fall in government spending, and sharp declines in consumer and investment demand led to a recession With similar policies enacted by Thailand's neighbors, the downturn soon spread across the region.

The speed and severity of the East Asian currency and financial crises took both investors and economists by surprise. Even economists who were not bullish about Asia in the long term, such as Paul Krugman (1998), admit that the crises caught them off guard. Debate is heated about the origins of the crises and about the best policy response by individual governments and the international community.

The possible causes of any financial crisis fall into three categories:

- Exogenous shocks or developments in international markets
- Economic mismanagement on the part of the crisis countries
- Intrinsic instability in international financial flows, which presumably results from failures in the international financial architecture governing those flows.

Because no single element is likely to have caused the East Asian crisis, the issue is the degree to which each of these different factors contributed to its onset and severity.

One of the interesting things about the East Asian crisis is that it was not precipitated by an exogenous shock. The global economy had been growing reasonably well, with stable commodity prices and low interest rates. Several specific developments, however, did affect the East Asian countries now in crisis. China had emerged as a major export competitor, and in the wake of the North American Free Trade Agreement (NAFTA), Mexico had become a formidable competitor as well. Furthermore, in the run-up to the crisis, the U.S. dollar gradually appreciated against the yen, which disrupted the exchange rates of countries, such as Indonesia and Thailand, that were de facto pegged to the dollar. But none of these developments can qualify as a major shock, a fact that shifts the focus to the other two potential causes.

With some risk of caricature, we can refer to these as the "bad policy" and "bad luck" hypotheses. The "bad policy" view is that countries such as Thailand made major mistakes in macroeconomic management and—more important—in the governance of the financial and corporate sectors. The policy regime encouraged borrowing abroad to invest in what we call the "wrong capital stock"; this borrowing was bound to lead to major problems sooner or later.

The "bad luck" view holds that the East Asian countries were the victims of a shift in investor expectations that became self-fulfilling but that did not have to happen. As Steven Radelet and Jeffrey Sachs (1998: 3) put it:

> To be sure, there were significant underlying problems besetting the Asian economies, at both a macroeconomic and microeconomic level (especially within the financial sector). But these imbalances were not severe enough to warrant a financial crisis of the magnitude that took place in the second half of 1997. . . . A combination of panic on the part of the international investment community, policy mistakes at the outset of the crisis, and poorly designed international rescue programs have led to a much deeper fall in (otherwise viable) output than was either necessary or inevitable.

This view emphasizes weaknesses in international financial architecture (no lender of last resort or bankruptcy proceedings to bring about orderly workouts). It also suggests that for individual countries such as Thailand, it was chance or bad luck that led to a financial run and brought them down.

Although the origins of the crisis have been hotly debated, analysis of the microeconomic foundations of the crisis has been hampered by the lack of systematic, high-quality data at the level of the firms or industries involved. From October 1997 to March 1998, working with the Ministry of Industry in Thailand, we conducted a survey of 1,200 randomly chosen plants in five tradable goods industries—auto parts, electronics, food products, garments, and textiles—to gather relevant production and financial information for the preceding three years. A range of questions about the existing situation, borrowing practices, major bottlenecks to increased competitiveness, human resources, and so on was also included. (Details of the sample are provided in the appendix.) These data can be used to examine some of the specific hypotheses that have been offered to explain the onset of the crisis, including

- The withdrawal of credit from viable firms (because of developments in the international capital market or other exogenous changes)
- A mismatch of short-term borrowing and the maturation of real investments financed by credit
- Failure of manufacturing firms that borrowed abroad to hedge their risk
- Continuing high levels of investment by firms despite declining profits.

Our findings support the view that poor policy is an important factor explaining the crisis. Specifically, policy incentives induced firms to take risky financial positions, financing long-term investment with large amounts of short-term debt. Firms that borrowed in foreign currency did not hedge their positions. Even before devaluation and the start of the recession in mid-1997, financial profitability had dropped sharply in these tradable goods industries. Thai firms were earning poor returns from their investments before international investors lost confidence in Thailand, even though

investment continued at a high level. Because only 40 percent of firms provided audited statements to their banks, the financial system—which had borrowed unhedged abroad—was slow to understand and respond to the problem of inefficient investment.

We also look at long-term structural issues of concern to firms. The interrelated issues of customs administration, red tape, and corruption are perceived as major bottlenecks to increased production and exports, more important than weaknesses in the financial system. Thailand also has well-known human resource problems—a shortage of technical and engineering skills and the high relative pay of production workers—and these are confirmed by this survey.

The data also show that the initial adjustment process undertaken in the fall of 1997 did not work smoothly. The five tradable goods industries that we surveyed should benefit from the standard adjustment package of fiscal and monetary austerity plus devaluation. In fact, capacity utilization and employment fell in all five sectors between the first half of 1997 and the second half. Domestic demand dropped sharply—a critical development for the textile and auto parts industries, which are domestically oriented. Auto parts manufacturing is the classic example of investing in the wrong capital stock: domestic demand is likely to be depressed for years, and shifting to exports will be difficult, given the cartelized nature of the industry and worldwide overcapacity. Even large export firms in electronics, food products, and garments were in trouble in Thailand because of a drop in foreign demand. Interestingly, the survey uncovered little evidence of a classic credit crunch—firms with orders but unable to obtain working capital.

Thus, the main story is that the combination of investment in tradable goods sectors oriented primarily to the domestic economy, the drop in regional demand, and bureaucratic bottlenecks (such as inefficient customs administration) made it difficult for Thailand to increase exports quickly. Hence the brunt of adjustment had to fall on the compression of domestic demand and imports. In light of this weak export potential, the initial adjustment program was probably too tight and caused an unnecessarily deep recession. Once the severity of the decline (as measured in part by these data) was evident, the austerity policies were moderated. In the spring of 1998 the Thai government started to relax its previous goal of achieving a 2 percent budget surplus.

In summary, there is ample evidence that Thailand made policy mistakes and that these mistakes created a vulnerable environment in which even mild changes in the international market could set off a serious downturn. At the same time, it is hard to understand the severity of the crisis without appealing to failures in the international architecture governing capital flows. In other words, bad policy ignited Thailand's crisis, but it was bad luck to have to face a world economy that was ill equipped to limit investor panic or to provide a framework for orderly workouts between international creditors and debtors. Looking to the future, the combination of a more

stimulative macroeconomic stance and a credible program for addressing structural weaknesses is the best hope for restoring confidence in Thailand, although the country will find it difficult to prosper as long as the region is suffering.

Macroeconomic Context

Although no one predicted the depth of the East Asian crisis, an examination of the macroeconomic variables points to several signs, evident before July 1997, that the existing situation was unsustainable. A correction was in order, even if not of the magnitude that has since unfolded. This is particularly true of Thailand, the country held responsible for setting off the regional crisis by devaluing the baht. The case for self-fulfilling panics is strongest for Indonesia, but even for that country, the evidence clearly shows that poor policy management and weak supervision and regulation contributed to the fragile position of the corporate and financial sectors.

Relative prices and interest rates affect the decisions firms make, particularly about investment allocation, financial structure, and the markets to target. The effect is clear in the aggregate data and is reinforced by findings from the newly available microdata.

As in neighboring countries, the Thai government was committed to a monetary policy that was inconsistent with domestic developments and the state of the financial sector. Officially, the baht was fixed to a basket of currencies, but in practice it was pegged to the U.S. dollar. Despite a burgeoning current account deficit, the Bank of Thailand repeatedly stated its commitment to the official rate of exchange. This exchange rate policy had two primary effects. The first was a loss in international competitiveness that stemmed from the choice of peg. The second was heightened borrowing from abroad that was encouraged by the implicit guarantee of an exchange rate parity.

Pegging the baht to the dollar had proved beneficial in earlier years because the dollar had depreciated relative to the yen and the deutsche mark, which made Thai exports attractive. But between 1994 and 1996 the dollar appreciated almost 40 percent relative to the yen. With Japan as one of the primary destinations for their goods, Thai firms found their exports declining sharply—not only to Japan, but also to third markets in which they competed with Japanese producers. A further effect was that direct investments from Japan slowed down, depriving countries of the long-term investments that are crucial sources of innovation and productivity growth.

With the exchange rate assumed to be fixed, firms took advantage of the lower interest rates on offshore loans and significantly increased their borrowing from such sources. This option was greatly helped in 1993 by the establishment of Bangkok International Banking Facilities (BIBFs), which loosened regulations on foreign borrowing by Thai banks. Believing that the exchange rate regime would be maintained, firms undertook large, unhedged positions.

With the nominal exchange rate fixed and domestic prices climbing, the rise the price of tradables relative to nontradables, while not unprecedented, was s considerable. Between 1994 and mid-1997, according to IMF statistics, the real a preciation was about 20 percent; as a result, export growth, which had been close 20 percent, turned mildly negative in 1996 and fell in the first half of 1997. Th current account deficit reached 8 percent in 1996.

The monetary authorities made some attempt to sterilize capital inflows as a mean of limiting the growth of domestic credit, but the rise in domestic interest rates together with the commitment to the baht-dollar exchange rate, made offshore bor rowing even more attractive. In 1996 almost two-thirds of all private capital flows to developing countries went to East Asia. As a share of the receiving country's gross domestic product (GDP), those inflows were by far the largest for Thailand, equaling 15 percent of GDP between 1994 and 1996. Private capital flows accounted for 5 percent of GDP in the Republic of Korea and in Indonesia and approximately 8–10 percent in Malaysia and the Philippines. In Thailand the net foreign liabilities of the financial institutions rose from 6 percent of deposits to one-third of deposits by 1996 (World Bank 1998).

The increase literally swamped domestic financial institutions with funds. The quality of the intermediation of these funds was critical. With banks providing the bulk of all financial services, the ability to supervise and regulate the banks was the key to determining how the funds were allocated. In practice, neither Thai firms nor Thai banks had much experience with such volumes of capital, and there was scant supervision. Consequently, much of this money was channeled into risky projects and highly cyclical sectors. Nonproductive investment in sectors such as real estate expanded dramatically. In Bangkok the amount of new office space quadrupled between 1994 and 1997. The boom in domestic assets reinforced an optimistic investment cycle, particularly since real estate was often used as collateral for further borrowing. The growth rate of loans is striking: close to 25 percent in 1996, the second-highest rate in Asia and more than double the growth rate of nominal GDP (World Bank 1998).

Thailand's experience demonstrates that high rates of investment are not sufficient to sustain a large current account deficit and an expansion of domestic credit. Although most of the new funds were invested rather than consumed, the investment was often misdirected. The concentration of resources in real estate has received considerable attention, but in fact, large amounts of investment went to industry. Here, too, many of these funds appear to have been allocated inefficiently. Few investments were directed to activities that would earn foreign exchange; rather, they were used to update facilities for domestic production or to buy low-return assets. These choices, along with real exchange rate appreciation, were associated with the sharp decline in financial profitability. The accumulation of the wrong capital stock still has consequences for the ability of firms to increase exports now that the asset bubble has burst and the baht has depreciated so drastically.

Microeconomic Foundations of the Crisis

At the microeconomic level, the two main points to emerge from the survey data are, first, that industrial firms continued to invest at a high level in 1996 and the first half of 1997 despite a sharp drop in sales, exports, and profitability, and, second, that much of this investment was financed by short-term borrowing that was not backed by collateral or by audited financial statements. This pattern of corporate and financial management left the country vulnerable to even small external shocks.

The five sectors covered by the survey differ in the extent of their export orientation, as seen in table 1. With the exception of the textile sector, 1995 was a growth year for both exporters and nonexporters. This was the heyday of the Thai boom. In sharp contrast, 1996 was a year of small increases in output (electronics, garments, and textiles) or decline (auto parts and food products). Hardest hit were medium-size exporters (those that shipped between 5 and 50 percent of their total output abroad), which declined more than 10 percent from their 1995 levels. Exports of domestic firms fell more than 20 percent, a sharp erosion of the competitiveness of firms without foreign linkages.

The survey also provides information on primary export markets, which differ considerably from sector to sector (table 2). The garment industry is focused on Europe and the United States (whose economies have continued to hum along fairly well), whereas the electronics industry exports to Japan and other Asian economies. Although the auto parts industry is not an overwhelming exporter, it is linked almost exclusively to Asian countries.

In this kind of survey, the information on profits is likely to be among the most problematic, but it should be possible to pick up broad trends. The most striking feature is the decline in profits for four of the five sectors in 1996; food processing was the exception, with its losses declining (table 3). Most of the downturn was concentrated in domestic firms and in firms that export less than half their output. These findings are consistent with the story that emerged from the output data—that domestic firms had lost international competitiveness.

While output and profitability were declining, firms continued to expand employment and—most strikingly—investment. Employment in firms with direct foreign investment surged between 1994 and 1995, then leveled off to a steady 3–5 percent increase in 1995 and 1996 (table 4). The growth of the capital stock was much more rapid. Such growth can be measured either as changes in total reported assets or as reported investment relative to the previous year's capital stock (and, fortunately, the two approaches provide a consistent story). The growth rate of assets of domestic firms did slow, from 22.5 percent between 1994 and 1995 to 14.9 percent between 1995 and 1996. Even so, this latter figure is particularly striking because it shows a high rate of growth of the capital stock during a year in which output and exports declined and profits were negative. The measure of investment

Table 1. *Output and Growth Performance, Thailand*
(percent, unless otherwise noted)

Sector	*Fiscal year*	*Output (baht)*	*Growth*	*Growth of domestic sales*	*Exports*	*Share of output exported*
Garments	1994	28,363,202	—	—	—	73 3
	1995	32,985,333	16.3	–1.2	22.7	77.3
	1996	34,385,143	4.2	7 2	3.4	76.7
Textiles	1994	48,784,260	—	—	—	33 4
	1995	43,865,237	–10 1	–8 3	–13 7	32 1
	1996	44,637,853	1.8	–3.7	13 4	35 8
Electronics	1994	155,699,056	—	—	—	72.6
	1995	204,569,102	31.4	31.0	31.5	72.7
	1996	210,347,504	2.8	6 0	1 6	71 9
Food processing	1994	42,072,321	—	—	—	62 6
	1995	63,820,112	51 7	31.5	63 8	67 5
	1996	50,032,865	–21.6	29.8	–46 3	46.2
Auto parts	1994	27,340,674	—	—	—	13.3
	1995	34,584,003	26.5	29 4	7 6	11 3
	1996	34,004,553	–1.7	–3.4	11.6	12.8
FDI status						
Domestic firms	1994	121,279,849	—	—	—	47.9
	1995	153,585,235	26 6	11 3	43.3	54.2
	1996	137,995,920	–10 2	3 9	–22 0	47.0
Foreign invested	1994	180,956,784	—	—	—	67 5
	1995	226,213,890	25 0	25.8	24.6	67 3
	1996	235,380,778	4 1	7 0	2 6	66 3
Export status						
Nonexporters	1994	60,385,825	—	—	—	0.4
	1995	70,198,695	16 3	16 1	n.a.	0 6
	1996	72,637,061	3 5	2 8	n a	1.2
Median exporters	1994	65,005,821	—	—	—	25 5
	1995	92,621,358	42 5	19 1	110.8	37 7
	1996	82,158,800	–11.3	2 4	–33 9	28.1
Large exporters	1994	176,867,866	—	—	—	92.3
	1995	217,003,733	22 7	25.3	22 5	92.2
	1996	218,612,058	0 7	27 0	–1.5	90.1
Total (942 firms)	1994	302,259,512	—	—	—	59.6
	1995	379,823,786	25.7	18 3	30.7	62 0
	1996	373,407,918	–1 7	5.5	–6.1	59 2

— Not available

n a. Not applicable.

Source. Thai Competitiveness Survey

Table 2. *Largest Export Markets, Selected Sectors, Thailand*
(percent)

Sector	*Asia*	*United States*	*Europe*	*Latin America*	*Africa*	*Other*	*Total firms*
Auto parts	64.0	16.0	11.0	1.0	4.0	4.0	96.0
Electronics	64 0	21.0	11.0	0 0	3.0	1.0	142.0
Food processing	74.0	16.0	5 0	0.0	4 0	1.0	80 0
Garments	23.0	35 0	35.0	0 0	5.0	1.0	202.0
Textiles	41.0	28 0	22 0	1 0	6.0	2 0	108 0
Total number of firms	301.0	158.0	125 0	3 0	29.0	12.0	628.0
Percent	48 0	25 0	20.0	0.0	5 0	2.0	100.0

Source Thai Competitiveness Survey

Table 3. *Net Profit as Share of Total Output, Selected Sectors, Thailand*
(percent)

Sector (number of firms)	*1994*	*1995*	*1996*
Auto parts (89)	6 17	5 69	5 13
Electronics (134)	4.26	3.78	2.45
Food processing (46)	0 46	−1.02	−0.45
Garments (208)	1 55	2.37	1 04
Textiles (157)	1 91	1 64	−0 43
FDI status			
Domestic (490)	1 67	1 63	−2 38
Foreign invested (144)	3 99	3 53	3.10
Export status			
Nonexporters (363)	3 57	3 12	−2 39
Median exporters (79)	6 15	5.62	4.25
Large exporters (192)	1.50	1 72	1.41
Total (634)	3.24	2 96	1 44

Source Thai Competitiveness Survey

relative to assets was even higher—a change of 18.5 percent between 1995 and 1996 for domestically owned firms and a similar 18.4 percent for those with foreign investors. About two-thirds of the total investment was in machinery and equipment, and one-third was in plants and land (unfortunately we cannot determine the share that was devoted to real estate investment). However, even non–real estate investments—including those financed by borrowing abroad—are not necessarily profitable. Although the data show clear signs of trouble by 1996, investment continued to expand at a high rate.

Table 4. *Measures of Expansion, Thailand*

Sector	*Change in employment*	*Change in total assets*	*Total investment as a percentage of total assets*	*Investment in machinery as a percentage of total assets*
Domestic				
1995	5.4	22.5	24.0	16.7
1996	3.5	14 9	18.5	11.3
Foreign invested				
1995	27 8	21 1	21 8	15.7
1996	3 7	17.7	18 4	13.6
Total				
1995	11 0	22 1	23.3	16.4
1996	3 6	15 7	18.5	12.0

Source Thai Competitiveness Study

How Did Firms Finance This Investment?

The figures confirm that the financial positions of Thai firms were particularly vulnerable. Firms were highly leveraged and extremely reliant on short-term financing. The mismatch of the maturity of their liabilities and their investments meant that changes in interest rates and the availability of credit were likely to have powerful effects.

At the end of 1996 debt-equity ratios were close to 2.5, on average, and in some cases above 5 (table 5).[1] About 35 percent of the firms had debt-equity ratios above 3—substantially lower than in Korea or Japan, but still much higher than emerging-market standards. Thus the Bank of Thailand's decision to raise interest rates to try to contain the depreciation of the baht had serious implications. The debt-equity ratios show that many firms had indeed borrowed heavily but since the onset of the crisis were acting to reduce the extent of their liabilities.

Among the five sectors, the textile industry had the highest debt-equity ratios and food processing the lowest. Setting aside the currency effect, the auto parts industry was the only sector that increased its debt-equity ratio after the crisis. Because this sector had the largest number of firms with long-run financing, shielding them from the fluctuations in interest rates, auto parts firms were more reluctant than other sectors to lower their debt positions, and it is likely that the value of their equity has fallen the most dramatically due to the utter collapse in their domestic demand.

Not surprisingly, large firms and exporting firms had higher debt-equity ratios than small firms and nonexporters (see table 5). Joint ventures, however, had lower ratios than local firms, a fact that is consistent with their ability to gain access to

Table 5. *Profile of Manufacturing Firms, Thailand*

Firm characteristics	*Size*			*Export orientation*			*Firms receiving foreign direct investment*			*Total*
	Small	*Medium*	*Large*	*None*	*Some*	*Major*	*Local*	*Some*	*Major*	
Financial indicators										
Short-term debt/total financing (percent)	84.0	81.0	80.0	83.0	79.0	82.0	8.0	85.0	77.0	81.0
Median debt-equity ratio	1.2	1.9	2.1	1.7	1.7	2.0	2.0	1.4	1.7	1.8
Mean debt-equity ratio	2.0	2.5	2.7	2.4	2.2	2.5	2.6	1.9	2.3	2.4
Share with audited statements (percent)	22.0	36.0	62.0	32.0	57.0	48.0	35.0	54.0	48.0	39.0
Share offering collateral (percent)	52.0	65.0	59.0	57.0	58.0	63.0	36.0	47.0	70.0	59.0
Firm composition										
Number of employees	43.0	123.0	796.0	132.0	491.0	538.0	202.0	435.0	643.0	288.0
Share that export (percent)	30.0	58.0	86.0				47.0	78.0	96.0	57.0
Share foreign invested (percent)	13.0	25.0	48.0	14.0	45.0	48.0	n.a.	n.a.	n.a.	28.0
Response to the crisis (percent)										
Current capacity utilization	56.0	65.0	71.0	59.0	65.0	73.0	62.0	66.0	76.0	64.0
Share with few workers	60.0	51.0	52.0	62.0	59.0	42.0	57.0	58.0	39.0	55.0
Optimistic about future growth	13.0	24.0	29.0	15.0	22.0	35.0	19.0	25.0	35.0	22.0
Total	390.0	396.0	362.0	648.0	141.0	342.0	882.0	223.0	122.0	1,227.0

n.a. Not applicable.

Source: Thai Competitiveness Survey

equity finance through their foreign parent or partner firms. Firms in the province had mildly higher debt-equity ratios than those in Bangkok.

With the onset of the crisis, debt-equity ratios fell across all types of firm. Large firms experienced the greatest decline; currently the degree of leverage differs little by firm size. Of course, the exchange rate effect significantly increases the ratios of firms with foreign currency liabilities, but as discussed below, at the time of the survey it was not readily apparent that these firms were reducing their foreign liabilities.

The more recent the information on debt-equity positions, the stronger the trend in the decline in the ratio. This could be consistent with a credit crunch, as loans are not being rolled over and the firms are carrying less debt. But it is also consistent with firms cutting back on their use of credit in the face of falling domestic demand. (It may also be due to some sample selection bias; that is, the most indebted firms were already bankrupt or were reluctant to participate in the study.)

Borrowing in Foreign Currency

With the dramatic depreciation of the baht, the costs of debt servicing for firms with extensive debt denominated in foreign currencies soared. The results of our survey, however, suggest that the problem is not as extensive in the manufacturing sector as might be feared. The large majority of firms borrowed only in baht, and those that did borrow in foreign currency were generally the more efficient firms. The issue is more severe in the private banking sector, particularly in the case of financial institutions that borrowed offshore and then lent domestically, assuming the exchange rate risk themselves. Of the 1,200 firms surveyed, only 140—predominantly large exporting firms with ties to foreign companies—reported that their liabilities were denominated in foreign currencies. Three-fourths of these reported that their foreign currency debt was less than the value of their baht-denominated debt at the end of 1996; half of the 140 said their foreign currency liabilities were less than 50 percent of the value of their baht liabilities in 1996. Interestingly, the maturities of the different liabilities are not highly correlated. Firms with foreign currency liabilities were less reliant on short-term loans and their short- and longer-term liabilities tended to be in different currencies. It is true that firms with foreign currency debt were generally more leveraged; firms with low debt-equity ratios rarely had any foreign currency liabilities.

Debt denominated in foreign currency does not need to be an issue unless firms do not hedge their positions. Because they assumed that the pegged exchange rate would remain in effect, however, only 19 percent of Thai firms with substantial foreign-currency-denominated debt had hedged some of their debt, and only 12 firms had fully hedged their positions. There is some natural hedging if firms export their goods, but most of the firms that borrowed in foreign currency did not export the majority of their goods. These statistics do not change when comparing firms' positions at the end of 1996 and at the time of the interviews in late 1997 or early 1998.

Equally worrying is that most firms had not altered their exposure to currency fluctuations in the months after the initial depreciation of the baht. Since most of the foreign currency debt was due within a year, these firms faced substantially higher financing costs in baht. Although a heavy reliance on short-term finance compounds the difficulties for firms with foreign liabilities, the relative lack of long-term finance is a more pervasive problem.

Reliance on Short-Term Financing

Leveraged debt or exposure to exchange rate fluctuations is less of a problem if the changes are likely to be temporary and if most of the debt matures over a period of several years. Thai firms, however, were not only highly indebted but also had borrowed in predominantly short-term instruments. Thus, the effects of the depreciation and the rise in interest rates were felt all the more keenly. This problem is not unique to the East Asian crisis, nor is Thailand's heavy reliance on short-term financing unusual for the region. But compared with other middle- and high-income countries, its firms have some of the lowest ratios of long-term debt to total assets. On average, almost 80 percent of liabilities are short term. In 1996, 56 percent of survey respondents said that they relied exclusively on short-term financing; by 1997 the number had risen to 59 percent. No clear patterns of characteristics identify firms with greater access to longer-term financing. What can be said is that before the crisis, there was little difference in the reliance on short-term financing across firm size or export status (although large exporting firms were slightly more likely to depend exclusively on short-term finance), but after July 1997 only large firms reduced their share of short-term debt.

An examination of debt-equity ratios together with the maturity of liabilities shows that before the crisis, firms with low debt-equity ratios tended to rely on short-term financing. One explanation is that these firms would not qualify for much financing. Conversely, firms with less reliance on short-term financing tend to have high debt-equity ratios and greater access to capital on more favorable terms.

The crisis has reinforced the case for improved access to longer-run finance. A better matching of liabilities and asset maturities would shield firms from much of the impact of monetary shocks. Another lesson is that while longer-term finance is desirable, firms must demonstrate that they qualify for it. Thus, improvement in issues of corporate governance and transparency must go hand in hand with financial reform.

Transparency and Disclosure

A main criticism of the East Asian financial systems is that there is insufficient transparency and disclosure of financial information, making it difficult to assess the risk

for potential borrowers. This is true not only for the financial sector but also for the manufacturing sector. To get at the issue of transparency and disclosure, the survey asked firms whether they had to provide audited statements to qualify for a bank loan and whether the loan required collateral. Only 40 percent said they were required to provide audited account statements to receive bank credit. There is little correlation between the need for audited statements and reliance on short-term credit, so many longer-run loans were offered with little objective disclosure of the firm's position. At the same time, those without audited statements were much more likely to have very low debt-equity ratios, and vice versa, which would be consistent with the disbursement of only small loans to firms whose books are not audited.

Sixty percent of respondents, most of them foreign-owned firms, said that their loans were backed by collateral. Many firms with short-term debt used collateral, firms with little short-term debt did not. Sixty-five percent of firms with high debt-equity ratios provided collateral, as did 72 percent of those providing audited statements but only half of those without audited statements.

Collateral offers some insurance to the lenders, but many firms with viable business plans have insufficient collateral to qualify for the necessary financing. Increasing the use of audited statements would give financial institutions a better view of the creditworthiness of an operation and a more reliable way to assess the true risks involved.

Response to the Adjustment Program

Thailand got into trouble because it had an incentive regime that encouraged firms to take out short-term loans in foreign currency to finance long-term investments—many of which were aimed at the domestic market. This system was inherently risky, and weak accounting and transparency rules made it even more difficult for financial institutions to assess the true risks involved. When foreign lenders began to lose confidence in the economy and devaluation became necessary, Thai financial institutions and firms sustained large losses.

The withdrawal of foreign capital required an adjustment of the current account, and Thai authorities adopted a standard structural adjustment program with support from the IMF. The objective of the program was to compress domestic demand while limiting any fall in output; that is, to reduce imports and increase exports. The data show how well the adjustment process worked in its initial phase. The sectors covered by the survey—including the major export industries—are the sectors that should have benefited from real devaluation and increased their exports. If they did not, it is important to understand what bottlenecks were holding them back.

pacity Utilization

nost three-fifths of firms reported that they operated at a lower capacity at the ie of the survey (six to nine months after the adjustment was implemented in July)7) than they had at the end of 1996. More than 60 percent had already reduced ›duction in the first half of 1997 relative to 1996. This general pattern holds for sectors and is consistent with the evidence that export growth fell substantially :r 1995.

Some areas of the economy managed better than others. Large firms were operat- ; at a higher capacity than small firms and experienced significantly less of a drop- . The same was true for firms with links to foreign companies relative to wholly mestic firms, and for exporters versus those that produced solely for the local ırket. In the second half of 1997, there was a 17 percent gap in capacity utilization ·ween nonexporters and exporters, a 15 percent gap between small and large firms, d a 12 percent gap between local and foreign-linked firms.

Although before July 1997 more than 40 percent of firms had planned to expand, nost half of them had abandoned this path by the first quarter of 1998. Only 8 ·cent of the firms that in 1997 were not planning to expand decided to do so in 98. Only 17 percent of firms reported increased capacity utilization after July 1997; e-fourth also expanded in the first half of 1997. The large majority of these firms re exporters, but their increase was modest, averaging about 13 percent. Fewer than hird of them expected to increase their capacity utilization by the end of 1998.

Auto parts makers were the hardest hit, largely because of the domestic orienta- n of the industry, and were the most pessimistic about their prospects. A factory it gives a vivid picture of new plants full of machinery to produce for the local ırket—standing idle. By the beginning of November 1997, most large auto assem- ·rs had announced that they were suspending operations (or severely cutting back ›duction) for three to six months. One ray of hope came with the announcement 1998 that Toyota would be shifting some production to Thailand (to export to ›an) to take advantage of lower costs of production. The tariff and nontariff barri- in the auto industry, however, suggest that exports from this sector hold little ›mise of increased growth for the sector overall.

nployment

ıe human dimension of the decline in capacity utilization is reflected in the reduc- n in employment levels. Eighty percent of firms with lower capacity utilization had back on the number of employees. Thus, close to 60 percent of firms employed ver workers at the time of the survey in early 1998 than they did in July 1997. Very v firms reported firing workers; almost all the reduction came from attrition.

Exporters were only slightly less likely than nonexporters to reduce their employee rolls. Auto parts manufacturers, which tend to have small work forces, suffered the most; only one in four had the same number of employees at the time of the survey as before the onset of the crisis. The reduction in other sectors was less dramatic. In textiles and electronics, slightly more than 50 percent of firms had fewer workers than they did in July 1997, while the figure in garments and food was about 45 percent.

Looking at the capacity utilization and employment results together, it appears that export firms were only slightly better positioned than domestically oriented firms. Large exporters did not experience falls in capacity utilization, but neither did they increase production, and half of them employed fewer workers at the time of the survey than they did before July 1997. Finally, both small and large producers for the domestic market cut back on capacity utilization and employment.

Perceived Causes of the Slowdown

The causes for the slowdown, as reported by the firms, call into question the common explanation of a credit crunch. Sixty percent of firms said that a substantial decline in domestic demand and higher input costs as a result of a depreciated baht were the primary sources of difficulty. In fact, only a third of the firms cited access to capital as a major problem, although substantially more said the cost of capital was a problem. The cost of capital was brought up by two-thirds of the food and textile companies and by about half of the auto parts, electronics, and garment firms. The cost of capital was particularly hard on small firms and nonexporters, and it affected domestic firms more than foreign firms. Small firms, nonexporters, and firms without foreign ties put greater emphasis on the decline in domestic demand than did large firms, exporters, and foreign-owned firms. These results are entirely consistent with the responses regarding changes in capacity utilization and employment.

There is substantial variation across sectors in the perceived causes of the slowdown. Textile firms stressed the effects of the depreciation of the baht, and auto parts makers were most worried about the drop in domestic demand. Firms in the food processing industry were least concerned about the fall in demand, but a third still reported it to be a significant problem. Access to capital was virtually unmentioned in the electronics sector, but more than 40 percent of textile firms cited it as a problem.

Nonexporters were encountering greater difficulties from the crisis than were exporters—one piece of evidence supporting the potential for exporters to expand and lead the way out of the current hardships. Yet more than 40 percent of exporters, particularly in the electronics and garment industries, saw a drop in foreign demand as a significant difficulty. With half of firms' exports traditionally targeted within the region and with the surge in the supply of electronic components, the fall in foreign

demand was not unexpected, but it makes the traditional path to recovery a more illusory one.

Firms were also asked to rank the difficulties of access to finance—particularly in the longer run—on a scale of 1 to 5 (with 1 meaning no problem and 5 a severe problem). About 20 percent said it was a major problem, and 24 percent reported that it was no obstacle at all. Interestingly, it was more of an issue for large firms, particularly large food processors.

Comparing the results in the short run with those over a longer period of time, it appears that the immediate impact of the crisis on manufacturers did not primarily concern finance. As a longer-term development goal, however, particularly in an environment where firms are trying to expand, finance is still an issue that needs to be addressed. Certainly there is a heavy reliance on short-term borrowing. The average bank debt-to-equity ratio in the sample is more than 200 percent, and almost all of the debt is short term. The fact that interest rates have not risen too drastically may explain why access to finance is not identified as a first-order problem.

Financial Characteristics and Impact of the Crisis

Linking the information on the financial characteristics of firms to their responses to the crisis suggests that the firms that might have been considered vulnerable are in fact some of the sounder ones. That is, at the time of the survey, firms with high debt-equity ratios or foreign currency liabilities had maintained higher rates of capacity utilization and remained relatively more optimistic about the future. This underlines the importance of looking at characteristics such as size, export orientation, and links to foreign firms when assessing firms' overall vulnerability. It should also be noted, however, that these data come from responses gathered in the first six months of the crisis, and sustaining high levels of debt or foreign liabilities becomes increasingly difficult as time goes by.

At the same time, businesses with higher debt-equity ratios were more likely to find access to capital an issue (43 percent, compared with 21 percent of those with low debt-equity ratios), suggesting that financial institutions were reluctant to continue lending to highly indebted firms. Interestingly, access to capital was more likely to be an issue for firms that were required to put up collateral than for those with unsecured loans This is consistent with the banks' increasing demand for collateral to back loans.

Not surprisingly, the more leveraged firms were almost twice as likely to see the cost of capital as a bottleneck. Despite the depreciation of the baht, the cost of and access to capital were less problematic for those firms borrowing in foreign currency. The fall in domestic demand was also less of an issue for those firms, in part because many of the firms borrowing in foreign currency were large exporting firms with foreign partners.

Companies that were highly leveraged at the onset of the crisis were the most pessimistic about the prospects of increasing their level of output. Strikingly, firms that borrowed in foreign currency (generally, large joint ventures that export) were significantly more optimistic about the likelihood that their output would rise. Firms that relied purely on short-term financing were significantly more pessimistic: 45 percent said their output would decline, and only 11 percent thought it would rise. Concerns that were less reliant on short-term financing were more evenly split; close to a third said output would rise, and the same percentage thought it would decline.

An accurate measure of the degree to which there is a credit crunch in Thailand is difficult to obtain. It is not sufficient to show that interest rates are rising or that the amount of new lending is lower. Rather, one needs to show the extent to which profitable investments are not being financed that otherwise would have been. Falling debt-equity ratios and canceled plans for expansion are consistent with a credit crunch but do not prove that viable projects are not being funded. These responses are also entirely consistent with a fall in domestic demand for credit. And, except for large firms, the share of short-term debt went up in 1997 compared with 1996.

Some additional evidence does support a credit crunch. For one thing, the cost and reported difficulty of obtaining capital were rising during the time in which the interviews were held. Another potentially worrying sign was the increasing reliance on short-term finance by many of the firms one would expect to be well positioned to recover quickly from the impact of the crisis—exporters and joint ventures. This reinforces the point that one of the most important areas in which financial reform will aid firms is in increasing access of profitable firms to longer-run finance.

Bottlenecks to Long-Term Productivity Growth

In addition to assessing the impact of the crisis, the survey gathered information on the determinants of firms' competitiveness. That portion of the survey covered a range of questions about long-term productivity growth and competitiveness, including quantitative data on inputs, outputs, and value added, as well as qualitative information about various problems and bottlenecks. Although the input and output data are beyond the scope of this article, it is important to record firms' perceptions of the bottlenecks that retard their long-term development.

The survey asked respondents to rank the bottlenecks to production and productivity growth on a scale of 1 (no problem) to 5 (a severe problem). The exporting firms in the sample had three clear reactions. First, they expressed a serious concern about the interrelated issues of corruption, customs administration, and red tape. Second, they voiced a similarly large concern about labor market issues: labor costs in general, and the availability of technical labor in particular. Third, they raised no great concern about access to finance (not surprising for these international firms) or

infrastructure (moderately surprising, given the visible problems). As noted earlier, nonexporting firms said that finance was an issue, but overall they were less concerned about it than were exporters—a worrying sign for the prospects of an export-led recovery. There were some differences in perceived bottlenecks between foreign and domestic firms. Both groups worried about corruption, customs administration, red tape, and labor costs, but access to finance was more of a concern for domestic firms.

By sector, the electronics industry cited red tape, corruption, and the lack of technical workers as serious bottlenecks. Garment and textile firms were also troubled by labor costs, although garment firms were more concerned about customs administration and red tape. The electronics and garment industries were the strongest exporters. Auto parts and food products ranked labor costs and red tape as the two greatest problems. High labor costs were the common thread in all of the sectors. Considering that the survey was taken shortly after a major devaluation, this problem is a bit of a surprise.

Concerns about corruption are echoed in international surveys covering a wide range of countries. Compared with countries at the same level of income, corruption is high in Thailand. Firms also report a large degree of regulatory discretion, a characteristic that is highly associated with corruption (World Economic Forum 1998). Corruption in the customs administration is a particularly severe problem for exporters and other businesses that need imported machinery and inputs in order to participate in the international division of labor.

A smooth adjustment to the shock of the East Asian financial crisis required Thailand to compress demand while limiting the decline in output, something that could be accomplished only if exports increased. The real devaluation—which raised the relative price of tradables—should have spurred such an increase. Indeed, the structural adjustment program supported by the international financial institutions assumed a good export response

What went wrong? In sectors burdened with worldwide overcapacity (notably auto parts), there was little scope to use the vast capital stock that had been put in place. More generally, since half of all external sales go to other Asian economies, the regional recession made any increase in exports unlikely.

It is noteworthy that the survey found little evidence that lack of credit hampered firms, either in production or in exports. In fact, it found that red tape and poor customs administration were larger issues than the availability of finance. Nonetheless, reforms to raise accounting standards, require financial disclosure, and impose prudential regulation of financial institutions remain important so that firms can gain greater access to longer-run finance and improve the match of maturities between available loans and investment projects.

In retrospect, given the weak potential for increasing exports quickly, the macroeconomic program was too tight. Fiscal and monetary contraction led to a sharp

Table A-1. *Characteristics of the Sample of Firms, Thailand*

	Total	Average number of employees	Size			Volume of exports			Foreign direct investment	
			Small	*Medium*	*Large*	*Small*	*Medium*	*High*	*Yes*	*No*
Sector										
Garments	334	232	125	130	72	151	26	138	56	278
Textiles	306	275	102	98	83	189	41	47	60	246
Electronics	229	543	49	64	104	96	36	83	130	99
Food	122	465	26	32	51	56	8	47	35	87
Auto parts	236	233	88	72	52	156	30	27	64	172
Export status										
Small	648	132	304	220	106	648	n.a.	n.a.	93	555
Medium	141	491	19	39	72	n.a.	141	n.a.	63	78
Large	342	538	38	115	172	n.a	n a.	342	163	179
Total	1,227	288	390	396	362	648	141	342	345	882

n a Not applicable.

Source Thai Competitiveness Survey

decline in domestic demand. Because this drop in demand was not mitigated by an increase in exports, the decline in output was unnecessarily large. Once the extent of the decline was evident, fiscal and monetary policies were eased to provide more stimulus to the economy.

Appendix: Overview of the Competitiveness Study

The findings presented here are based on responses from more than 1,200 firms that were interviewed in the last quarter of 1997 and the first quarter of 1998. The detailed survey was designed to learn about firms' production, labor force training and turnover, technology acquisition, and financial structure. Questions sought information from 1994, 1995, and 1996, as well as from the period immediately following the onset of the crisis. These firms (employing a total of more than 350,000 workers) were drawn from five of the major tradable goods sectors, which should have been relatively well positioned to take advantage of the large depreciation of the baht. The five sectors selected, based on their shares in exports and GDP, were auto parts, electronics, food processing, garments, and textiles.

Each sector is well represented. The inclusion of fewer food processing firms reflects the smaller number of establishments in this sector (table A-1). More than half the firms surveyed export at least some of their output, although the share of firms that export varies from sector to sector. Electronics and garments have the most exporters: two-thirds of the firms export, while only 50 percent of auto parts and textile firms do. In this regard, the sample is representative of the industries more generally: electronics and garment firms are more export oriented, while auto parts and textile firms are more geared to the domestic market. This distinction is important in understanding the differences across sectors in the reaction to the crisis.

There is a good dispersion in the size of firms. Care was taken to include small and medium-size enterprises because they represent the bulk of firms and were most likely to have trouble coping with the recession. In this analysis, firms were divided into three categories based on total employment: small (under 50 employees), medium-size (50 to 150 employees), and large (more than 150 employees). The thresholds were chosen so that each category accounted for about a third of the total. Electronics and food processing have a noticeably higher proportion of large firms than do the other three sectors.

About 30 percent of the firms have a significant relationship with a foreign firm, either as a joint venture or as a wholly owned subsidiary of a foreign parent. Such relationships are most common among electronics firms and least common in garment firms. The tie to a foreign firm could be an important factor in giving firms access to capital or to markets overseas. Therefore it is not surprising that in the sample, close to 84 percent of firms with foreign ties export, compared with 51

percent of local firms. Part of the difference is also due to the fact that foreign firms tend to be larger. The average number of workers in a foreign exporting firm is 700, compared with 347 for local exporters; it is 185 for foreign nonexporters, compared with 83 for local nonexporters.

Notes

David Dollar is a research manager in the World Bank's Development Research Group Mary Hallward-Driemeier is an economist in the Development Research Group. They wish to thank Dennis Tao and Hairong Yu for excellent research assistance.

1 The survey gathered information on firms' balance sheets from before the crisis as well as after its onset The date of collection of the post–July 1997 data varies from the end of the third quarter of 1997 (when less impact might be expected) to the end of the first quarter of 1998. During this period, the baht-dollar exchange rate moved considerably, from 35 baht to the dollar, to 52, and back to nearly 40. This fluctuation raises the issue of the appropriate exchange rate to use in calculating the debt-equity ratios for firms with foreign liabilities In the analysis for 1997, the exchange rate of 36 5 prevalent at the end of the third quarter of 1997 is used.

Another solution would be to maintain the use of the pegged value of the baht before devaluation (25 baht per dollar) This would allow for a measure of the change in the "real" debt-equity position Clearly, the actual ratios are considerably higher for firms with foreign currency liabilities, but using an exchange rate of 25 does not alter the averages by very much because the number of firms with a significant share of their liabilities in foreign currency is small

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems

Krugman, Paul. 1998 "What Happened to Asia?" Massachusetts Institute of Technology, Cambridge, Mass Processed

Radelet, Steven, and Jeffrey Sachs 1998. *The Onset of the East Asian Financial Crisis* NBER Working Paper 6680 Cambridge, Mass. National Bureau of Economic Research

World Bank 1998 *East Asia The Road to Recovery* Washington, D C

World Economic Forum. 1998 *The Global Competitiveness Report* Geneva

Corporate Performance in the East Asian Financial Crisis

Stijn Claessens • Simeon Djankov • Lixin Colin Xu

The sharp decline in the once-stellar performance of East Asian corporations following the 1997 financial crisis has sparked an intense debate. Some observers argue that external shocks, including a drop in aggregate demand and a shortage of working capital, explain the corporate sector's poor performance. Others assert that the difficulties were apparent well before the crisis and that the risky financial policies pursued by these firms left them vulnerable. A survey of the literature shows little microeconomic evidence to support either view.

This article compares the growth and financing patterns of East Asian corporations in the years before the crisis with those in other countries. It finds little microeconomic evidence that corporate growth was weakening but some support for the argument that many firms had a weak financial structure that left them vulnerable to an economic downturn. Based on a sample of more than 850 publicly listed firms in the four crisis countries—Indonesia, Malaysia, the Republic of Korea, and Thailand—and two comparators, Hong Kong (China) and Singapore, it appears that firm-specific weaknesses already in existence before the crisis were important factors in the deteriorating performance of the corporate sector.

The East Asian crisis has sparked a large body of literature seeking to explain its causes, onset, and evolution. Whether sudden shifts in market expectations and confidence were the primary source of the financial turmoil has been hotly debated. Proponents of this view argue that although some macroeconomic and other fundamentals may have worsened in the mid-1990s, the extent and depth of the crisis can be attributed not to a deterioration in fundamentals but rather to the panicky reaction of anxious domestic and foreign investors (Furman and Stiglitz 1998; Radelet and Sachs 1998). Others argue that the crisis reflected structural and policy distortions in the region, including weak macroeconomic policies, and that fundamental imbalances triggered the crisis (Corsetti, Pesenti, and Roubini 1998).

The World Bank Research Observer, vol 15, no 1 (February 2000), pp 23–46

The contributions of macroeconomic weaknesses, moral hazard, regulatory deficiencies, and the inherent instability of financial markets have all been investigated in depth (see Kaminsky and Reinhart 1999 for a review of recent work). Only more recently have variables in the corporate sector itself—performance, financial structure, and corporate governance—been included as explanatory factors. Some observers have gone so far as to attribute the East Asian debacle to these factors (Harvey and Roper 1999; Johnson and others forthcoming; Krugman 1999). In hindsight, it has become apparent that the corporate financial structure of many companies was too weak to withstand the combined shocks of increased interest rates, devalued currencies, and sharp declines in domestic demand. Yet poor performance and risky financial policies were not notable features of East Asia's economic ascent before the financial crisis—quite the opposite. Considered by observers to be an important contributor to the East Asian miracle, the corporate sector was generally viewed as being very competitive and adept at exploiting new market opportunities. As a consequence it attracted considerable amounts of foreign capital. Hence, it is unclear whether the corporate structure compounded the financial crisis or whether corporations were the innocent victims of a financial crisis brought on by other factors.

In this article, we review the performance and financing patterns of East Asian corporations in the years immediately before the crisis. We analyze the return on assets, in real local currency and in dollars, and the debt burden of 5,500 firms in nine economies in the region and two comparator countries, the United States and Germany. We find that while performance was perhaps weakening in some East Asian countries, returns on assets were generally high; in many of the nine economies, returns averaged twice as high as those recorded in Germany and the United States. The high levels of investment dictated a heavy reliance on external financing, and because outside equity was used sparingly, debt levels were high in most of these economies and in fact were increasing in Malaysia, the Republic of Korea, and Thailand. Short-term borrowing became increasingly important, especially in Malaysia, Taiwan (China), and Thailand. Some of the vulnerabilities in corporate financial structures that are now seen as factors in triggering and aggravating the crisis were thus already in existence in the early 1990s.

We survey the nascent literature on the role of the corporate sector before and during the crisis as well as the interaction between characteristics of the corporate sector and developments in the general economy. One way to assess the relative importance of external financial shocks and underlying weaknesses would be to compare the operational performances of firms that had different exposures to these financial shocks. A dearth of data and the elapse of time since the crisis, however, make it difficult to separate out the causes and consequences. Much of the literature has been qualitative in nature, and very little of it has involved analyzing individual performance of corporations during or following the crisis.

We therefore provide a preliminary empirical assessment of the importance of various factors by reviewing the performance of a sample of publicly traded firms in four crisis economies, Indonesia, Korea, Malaysia, and Thailand, and two comparators, Hong Kong (China) and Singapore, before and during the East Asian crisis. We find that little can be explained by a firm's financial and other characteristics or by the institutional environment of the particular country. Idiosyncratic shocks appear to have been the most important factors in the sharp deterioration in corporate performance. Nevertheless, of the variance that can be explained, we find that firm-specific, nonfinancial characteristics were most important. Industry-specific shocks and the institutional environment also contributed to the decline in profitability, and financing patterns had a strong influence on operational performance.

Corporate Performance and Financing

The data used here are based on annual reports of the companies listed on the major stock exchanges and come from the Worldscope database (see Claessens, Djankov, and Lang forthcoming a). This data set is unbalanced; that is, the number of observations varies from year to year. We have excluded companies that reported fewer than three times between 1988 and 1996, financial and banking institutions, and companies that did not report all of the following variables: net sales, net income after taxes, total assets, and the value of common equity.

Our first measure of performance is the real rate of return on assets (ROA), in local currency, calculated as the earnings of a firm before interest and taxes divided by total assets, minus the annual inflation rate in the country. The advantage of this measure is that it is not influenced by the liability structure of the firm because it excludes interest payments, financial income, and other income and expenses. At the same time it is not a complete measure of firm productivity (unlike total factor productivity) because it does not control for inputs other than capital

East Asian corporations have had quite different ROAs, ranging from relatively low profitability in Hong Kong, Japan, Korea, and Singapore to high real ROAs averaging 9 to 10 percent in Indonesia, the Philippines, and Thailand (table 1). Returns in Malaysia and Taiwan fell between these two groups, but at about 7 percent, they were close to the high performers. All these ROAs were high compared with the returns of corporations listed on the DAX in Germany and on the New York Stock Exchange in the United States, which were about 5 percent The comparison indicates that the corporate sector's contribution to the East Asian miracle was significant during most of this period.

A further comparison of corporate performance in all the countries that report to Worldscope reveals that Thailand, the Philippines, and Indonesia posted the highest returns in this sample of 46 economies, followed by Taiwan and Malaysia (figure 1).

Table 1. ***Return on Assets in Real Local Currency, Selected Economies, 1988–96***
(percent, medians)

Economy	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1988–96*
Hong Kong (China)	5.1	5.3	4.9	4.8	4.5	3.8	3.9	3.9	4.1	4.6
Indonesia	—	—	9.4	9.1	8.6	7.9	7.4	6.2	6.5	7.1
Japan	5.7	5.4	4.6	4.7	4.8	4.5	4.1	3.8	3.6	4.1
Korea, Rep of	4.4	3.9	4.1	4.0	3.9	3.6	3.4	3.6	3.1	3.7
Malaysia	5.4	5.6	5.4	6.2	6.0	6.5	6.3	6.1	5.6	6.3
Philippines	—	—	—	7.1	6.4	8.1	8.5	6.8	8.4	7.9
Singapore	4.9	4.5	4.2	3.9	5.2	4.6	4.5	3.9	4.0	4.4
Taiwan (China)	—	—	—	5.1	6.2	6.5	6.8	6.5	6.6	6.7
Thailand	10.8	11.0	11.7	11.2	10.2	9.8	9.3	7.8	7.4	9.8
Germany	5.3	5.5	5.5	5.7	5.6	5.2	5.1	4.9	5.0	4.7
United States	4.7	4.8	5.1	4.9	5.2	5.4	5.3	5.2	5.2	5.3

— Not available.
Source Worldscope database

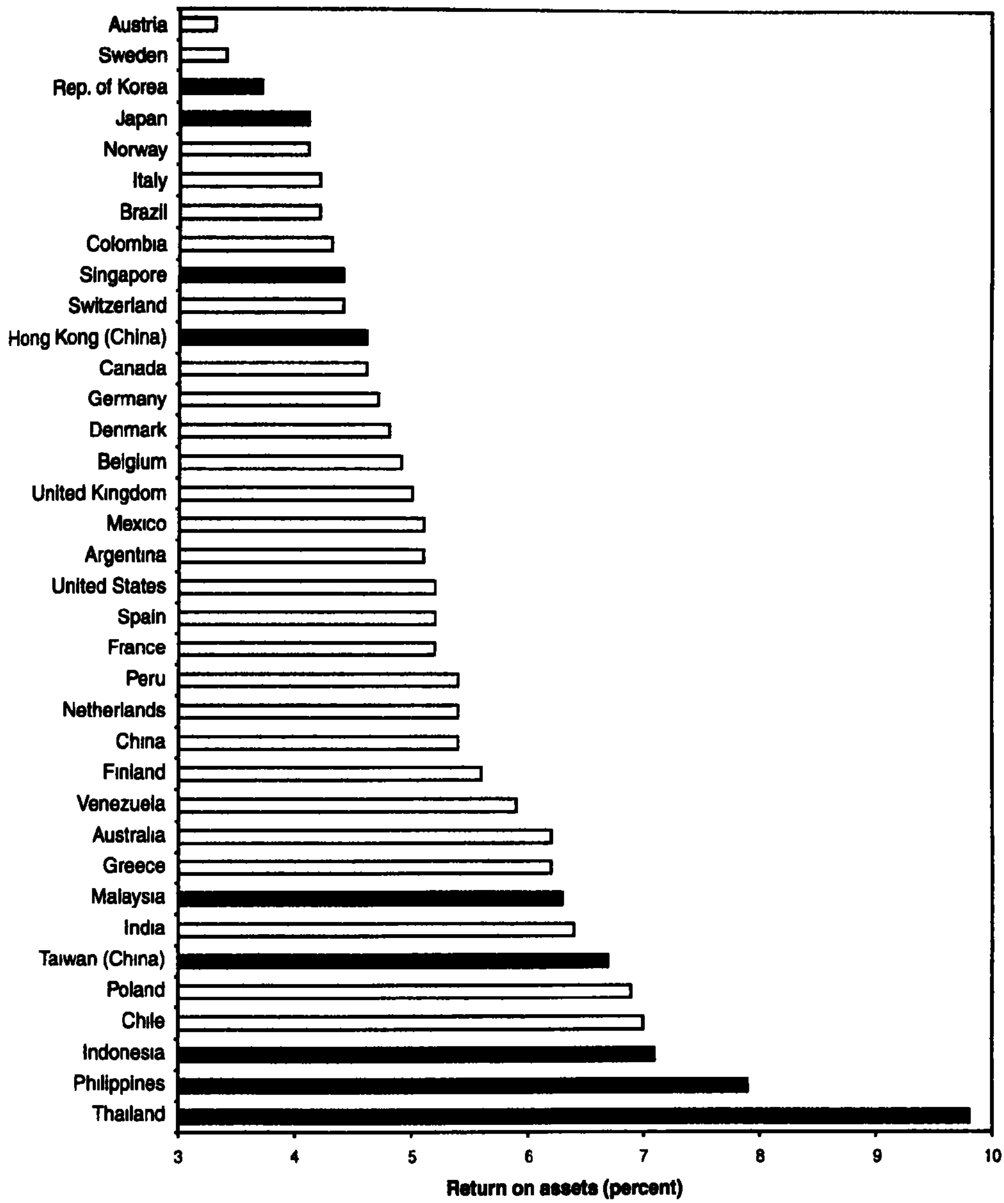
At the other end were Korea and Japan; Singapore and Hong Kong also had relatively low ROAs in real local currency.

Table 2 shows ROAs in U.S. dollars, adjusted for the effects of currency movements. This measure of performance represents the point of view of an international investor who can allocate resources across several countries. With the exceptions of Japan (6.6 percent) and Taiwan (8.4 percent), ROAs in all the East Asian countries were higher than the U.S. median (8.7 percent). The Philippines (18.7 percent), Thailand (14.7 percent), and Indonesia (13.0 percent) posted the highest average returns over the 1988–96 period.

An examination of the data shows the degree of risk inherent in the liability structures of East Asian corporations. High rates of investment meant that companies had to turn to external financing to make up for the lack of capital from retained earnings. In fact, firms in these countries have always relied on high levels of external financing, primarily from the banking system Leverage, measured by the ratio of total debt divided by equity, remained well above that in industrial countries (table 3). The most highly leveraged economy over this period was Korea; the lowest was Taiwan. Malaysia and Singapore were also low; leverage in the Philippines, while rising, was well below that of Indonesia and Thailand.

In the few years just before the crisis, leverage increased in Japan, Korea, Malaysia, and Thailand. Japan had reduced its leverage in the early 1990s, possibly in the course of a financial retrenchment, but when faced with a shortage of equity and other sectoral difficulties, companies there subsequently rolled over these loans. The rise in leverage in the Philippines probably resulted from reforms in the mid-1980s that revived the country's corporate and financial sectors and resulted in better financing possibilities.

Figure 1. *Return on Assets in Local Currency, Selected Economies, 1988–96*



Source Worldscope database, authors calculations

Table 2. *Return on Assets in Nominal U.S. Dollars, Selected Economies, 1988–96*
(percent, medians)

Economy	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1988–96*
Hong Kong (China)	8.0	8.4	7.2	12.9	14.3	12.5	11.5	8.0	10.3	10.3
Indonesia	—	—	16.0	13.7	12.6	15.3	11.7	10.7	11.2	13.0
Japan	6.5	–6.0	13.3	14.8	7.0	16.2	15.6	1.0	–9.2	6.6
Korea, Rep. of	25.1	10.3	7.3	7.2	6.4	5.9	12.1	9.9	–1.0	9.2
Malaysia	–0.8	8.8	7.2	9.9	14.8	6.1	15.5	12.2	9.5	9.2
Philippines	—	—	—	23.2	21.2	5.4	29.4	7.5	16.5	17.2
Singapore	8.9	9.4	15.6	13.6	6.9	9.3	16.4	9.0	6.8	10.7
Taiwan (China)	—	—	—	6.2	12.0	4.6	12.4	6.3	8.9	8.4
Thailand	13.9	14.6	19.3	16.9	13.4	13.1	16.6	13.2	11.5	14.7
United States	8.7	9.6	10.5	9.1	8.3	8.4	7.9	8.0	8.1	8.7

— Not available
Source Worldscope database

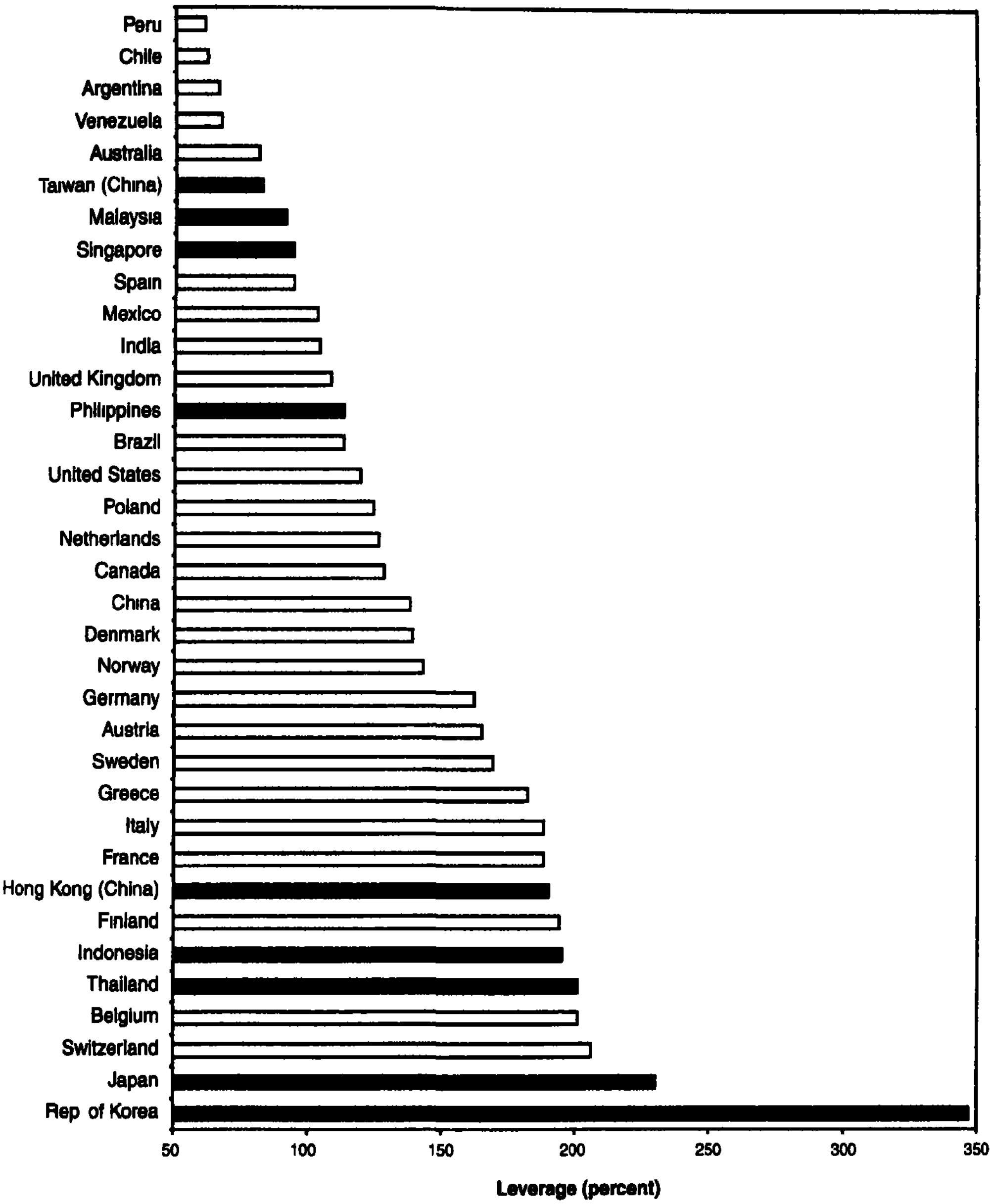
A comparison by country of the average leverage ratios of East Asian corporations from 1988 to 1996 shows that firms in Korea were the most highly leveraged, followed, in order, by companies in Japan, Thailand, Indonesia, and Hong Kong. At the opposite extreme, firms in Taiwan were the least leveraged, followed by Malaysia, Singapore, and the Philippines. The pattern across other regions is also interesting. Western European countries typically displayed high ratios of debt to assets, reflecting the bank-based nature of their financial systems, with Swiss firms almost as highly leveraged as Japanese firms. In contrast, corporations in Latin American coun-

Table 3. *Ratio of Debt to Equity, Selected Economies, 1988–96*
(percent, medians)

Economy	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1988–96*
Hong Kong (China)	1.832	2.311	1.783	2.047	1.835	1.758	2.273	1.980	1.559	1.902
Indonesia	—	—	—	1.943	2.097	2.054	1.661	2.115	1.878	1.951
Japan	2.994	2.843	2.871	2.029	2.042	2.057	2.193	2.367	2.374	2.302
Korea, Rep. of	2.820	2.644	3.105	3.221	3.373	3.636	3.530	3.776	3.545	3.467
Malaysia	0.727	0.810	1.010	0.610	0.627	0.704	0.991	1.103	1.176	0.908
Philippines	—	—	—	0.830	1.186	1.175	1.148	1.150	1.285	1.129
Singapore	0.765	0.922	0.939	0.887	0.856	1.102	0.862	1.037	1.049	0.936
Taiwan (China)	—	—	—	0.679	0.883	0.866	0.894	0.796	0.802	0.820
Thailand	1.602	1.905	2.159	2.010	1.837	1.914	2.126	2.224	2.361	2.008
Germany	1.535	1.552	1.582	1.594	1.507	1.534	1.512	1.485	1.472	1.514
United States	0.798	0.848	0.904	0.972	1.059	1.051	1.066	1.099	1.125	1.034

— Not available
Source Worldscope database

Figure 2. *International Comparison of Leverage, Selected Economies, 1988–96*



Source Worldscope database, authors' calculations

tries (Argentina, Chile, Colombia. Peru, and Venezuela) carried low debt ratios, reflecting the less-developed banking systems of these countries (figure 2).

Long-term debt as a share of total debt was low in East Asia during the entire period (1988–96), accounting for less than a third of all loans in Malaysia, Taiwan, and Thailand (table 4). Japan and the Philippines had the highest shares, at 48 and 52 percent, respectively, while the other economies in the region hovered around 43 percent. In contrast, long-term debt accounted for about three-quarters of total corporate debt in the United States and about 55 percent in Germany. Despite the close attention paid to the role of short-term debt in precipitating the financial crisis, the data do not suggest a massive buildup in such debt, at least up to the end of 1996. In fact, only Japan's share of long-term debt decreased during this period.

An international comparison of the maturity of debt structures shows that most companies in East Asian countries ranked below those in European and Latin American countries in their share of long-term debt.[1] Only in the Philippines was the average share of long-term debt more than 50 percent. In general, businesses in richer countries tend to have more long-term debt (Demirguc-Kunt and Maksimovic 1998). But many companies in higher-income East Asian countries rely less on long-term debt than would be expected on the basis of their per capita income level. Japan, for example, ranks below many other members of the Organisation for Economic Co-operation and Development (OECD). Whether this pattern of debt set the stage for the East Asian financial crisis is addressed in the next section.

The Role of the Corporate Sector in Financial Crises

Four hypotheses have been proposed to explain the role of corporate performance and financing patterns in triggering and aggravating the East Asian financial crisis We look at each in turn.

Hypothesis 1

The first view asserts that the weak corporate performance after the crisis was due largely to the aggregate shocks experienced by the East Asian economies, including declines in domestic and external demand, a withdrawal of (short-term) capital flows, a devaluation of currencies, and an increase in domestic interest rates (Furman and Stiglitz 1998).

Empirical work on the importance of aggregate shocks has been limited, although Dollar and Hallward-Driemeier, in an accompanying article in this issue, suggest that it played an important, yet not an exclusive, role in the deterioration of the corporate sector. Respondents to a survey of about 4,000 small and medium-size firms in late 1998 and early 1999, for example, said that the four most important

Table 4. *Long-Term Debt, Selected Economies, 1988–96*
(percent, medians)

Economy	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1988–96*
Hong Kong (China)	59.7	59.5	53.8	56.5	44.7	44.7	40.7	37.3	36.4	44.9
Indonesia	—	—	—	52.4	40.8	39.6	41.6	41.8	43.3	43.1
Japan	49.9	54.1	53.8	49.9	49.4	51.7	47.7	44.4	40.8	48.4
Korea, Rep. of	55.7	47.2	49.8	49.8	44.2	43.7	41.4	40.4	41.5	43.7
Malaysia	35.8	35.5	32.5	27.1	26.9	26.6	27.2	27.8	29.9	29.2
Philippines	—	—	—	57.2	53.1	50.3	50.2	49.8	51.4	52.2
Singapore	57.2	55.4	54.1	33.8	33.8	33.9	40.2	38.6	41.1	43.3
Taiwan (China)	—	—	—	53.9	44.4	32.8	34.6	34.3	38.9	35.9
Thailand	58.1	49.8	38.8	34.3	25.2	26.4	27.6	32.9	32.8	30.9
Germany	56.8	55.4	54.5	53.9	55.2	55.4	55.4	55.3	54.7	55.3
United States	77.7	77.2	76.3	76.7	75.8	76.2	75.2	74.6	74.1	75.9

— Not available

Source Worldscope database

causes of the decline in performance were a fall in domestic demand, an increase in the cost of inputs (caused by the currency depreciation), and a rise in interest rates and labor costs (Colaço, Hallward-Driemeier, and Dwor-Frecaut 1999). The surveys do not provide a quantitative measure of the importance of these shocks, however, nor do they permit an assessment of the degree to which they were aggregate exogenous shocks or episodes that were exogenous to the firm or to the country. It is possible, for example, that the most vulnerable countries were those whose firms were most exposed to international trade or were already highly leveraged at the outset of the crisis. It is also possible that the corporations in these countries took on riskier projects to cover the higher costs of borrowing.

One way to disentangle the relative importance of various shocks would be to explore the possibility that although the effects of these shocks are likely to differ by industry, they need not differ across firms in the same industry. In countries open to international trade, firms in the same industry producing tradable goods, for example, are likely to have been similarly affected by a decline in aggregate external demand. Because each country has a different industrial structure, these differences could explain some of the variation in countries' performance. So far, this type of analysis has not been conducted.

Hypothesis 2

The second view argues that the poor corporate performance after the crisis reflected prior fundamental weaknesses (Corsetti, Pesenti, and Roubini 1998). As noted, however, there was very little indication of declining profitability before the crisis. Ac-

counting measures do not adjust for risk, however, and the relatively high rates of return posted by firms in the region may have been inadequate for the degree of risk undertaken by East Asian corporations. Harvey and Roper (1999) find supporting evidence for this view. Using rates of return on stocks, they find that although capital markets in the region mobilized substantial amounts of new funds and enhanced their liquidity, risk-adjusted returns on the shares of the companies concerned were well below those generated in other equity markets in the 1990s, especially in industrial countries.

This line of reasoning assumes that before the crisis, weak firms operating with a high degree of risk were not being disciplined sufficiently through competition and monitoring by shareholders or creditors (foreign as well as domestic). Poor performers were not forced to adjust and raise their rates of return sufficiently to compensate investors for the risks taken. Instead, distressed firms were allowed to operate while their losses continued to grow. Anecdotally, there are many examples of firms that continued to function even though they were technically insolvent. For example, Alphatec, a Thai semiconductor maker, was declared bankrupt in 1996 but did not close down until 1999. In general, very few enterprises declared bankruptcy, suggesting that firms continued to borrow and that banks continued to overlook the rise in bad loans. Of course, there were many successful corporations as well, as the generally high rates of corporate growth attest, so the presence of some weaker corporations does not fully explain the systemic nature of the financial crisis.

This argument may also presume that profitability was overstated, in part because of the need to attract external financing. The lack of transparency may thus have postponed the crisis. A cross-country comparison by La Porta, Lopez-de-Silanes, and Shleifer (1998) suggests that relatively weak accounting standards in East Asian countries may have allowed firms in crisis-affected countries to shelter their actual financial position and continue in business even after they were no longer financially viable. Furthermore, creditors' rights were weakly enforced because the judicial systems in these countries were often inefficient. And some evidence supports the view that weak corporations relied excessively on new financing. In a cross-sectional sample of publicly listed corporations in Thailand, for example, an increase in leverage in 1996 over 1995 was correlated with declines in profitability—evidence that struggling firms relied on increased external financing to overcome declining earnings (Alba, Claessens, and Djankov 1998).

Several studies have shown that ownership structure may encourage a lack of discipline and induce risky behavior. In many East Asian countries, shared ownership and other links between banks and corporations were extensive and could have skewed the market's allocation of resources toward influential borrowers, introducing excessive risk. Claessens, Djankov, and Lang (forthcoming b) show that about two-thirds of the publicly listed corporations in East Asian countries belong to larger groups, many of which include one or more financial institutions. Such links reduce the disciplining

role that would otherwise be played by external financiers, who have an incentive to monitor their investments carefully. In Korea, for example, many conglomerates (called *chaebols*) had ownership links with nonbank financial institutions; the larger conglomerates were often linked with a main bank. Kim (1999) shows that Korean firms with ownership links to financial institutions were more highly leveraged and had more short-term debt than other firms. Ownership links also played a role in Indonesia, where many banks belonged to business groups, and in Thailand, where a small number of families owned both banks and corporations (Laeven 1999).

Governments also influenced the financial sector by owning banks directly and by granting preferential access to resources; government influence in turn injected a political dimension into lending decisions. In Indonesia, for example, about 50 percent of the banks were state-owned; in Korea, the government had traditionally been influential in the banking sector (Cho and Kim 1995). This influence may be associated with increased risk-taking and poor allocation of assets, weakening the financial sector in the process. Claessens and Glaessner (1997) note that an inadequate regulatory and supervisory framework allowed well-connected borrowers to become even more highly leveraged. Moreover, the process of financial liberalization, including capital account liberalization, may have made credit more readily available to the largest—but not necessarily the most efficient—firms.

Weak corporate governance may also have affected firms' behavior. Although a lack of shareholder rights is often cited as one cause of the crisis, the region was not out of step with other emerging economies in this respect. But equity protection was less effective than that in OECD countries; although creditors had the same protections in both sets of countries, what was missing in the East Asian economies was the enforcement of these rights (La Porta and others 1999). Moreover, measures commonly used to assess transparency and evaluate the environment for private business transactions show a high incidence of corruption in East Asian countries (Kaufmann, Kraay, and Zoido-Lobatón 1999).

Lack of transparency and weak governance bear major responsibility for the East Asian crisis, according to Johnson and others (forthcoming). Their model identifies a channel through which weak corporate governance results in more stealing by managers, which in turn leads to a sharp currency depreciation and ultimately to a recession. They find empirical support for their model in a sample of 25 developing countries. Other economists, however, reject this notion. Stiglitz and Bhattacharya (2000), for example, argue that increased transparency in the form of disclosure requirements is unnecessary because markets can and do provide optimal incentives for disclosure. Under certain circumstances, they note, disclosure of the information could actually exacerbate fluctuations in financial markets and precipitate a financial crisis. Furman and Stiglitz (1998) point to the fact that even countries with solid legal and regulatory systems and no transparency problems, such as Sweden, have had financial crises.

Hypothesis 3

A third strand of the literature relates to imperfections in financial markets. Aggregate and financial shocks affecting financial institutions can affect the corporate sector by curtailing credit, including working capital and trade financing, to borrowers with valuable trading and investment opportunities (see Kashyap and Stein 1994 for a review). Shocks, whether real (changes in the terms of trade), financial (increases in world interest rates and declines in external financing), or regulatory (increased capital adequacy requirements or tighter loan classifications), can cause a real or perceived shortage of capital for banks. As a result, banks may become unwilling to lend even to viable corporations and instead may prefer to invest excess liquidity in safe assets, such as government bonds. Such a decision will curtail the amount of financing available for investment (or even for working capital), which can impair firms' performance.

A credit crunch can originate from weak financial institutions or from tightened regulation and supervision. The latter has been analyzed extensively in the context of the Basle Accord of 1989, which tightened capital adequacy rules for some classes of internationally active banks. A review of the data (BIS 1999) suggests that the effects of these tighter rules on aggregate credit provision have been minimal, although there is some evidence that borrowers from weaker banks have been affected. Most of this work applies to industrial countries, however, where shocks have been small and where many alternative financing instruments have been available. Given East Asia's unbalanced financial systems, which are dominated by banks (Greenspan 1999), and the fragile state of some banking institutions even before the crisis, it is likely that, at least initially, East Asian corporations suffered a credit crunch induced by weaknesses in the banking sector and a tighter regulatory framework (Domaç and Ferri 1999).

Increased uncertainty about whether and at what price loans will be available can also result in a shortage of loanable funds (Stiglitz and Weiss 1981). These effects can be particularly severe for bank lending because banks are more likely than other financial intermediaries or markets to lend to firms that suffer from a greater degree of informational asymmetries. Thus small and medium-size firms are likely to be harder hit by shocks to the banking system or by a tightening of regulations and supervision. Furthermore, a so-called balance-sheet effect can further amplify the effect of shocks on corporations (see Bernanke and Gertler 1995 for a review). In the presence of informational asymmetries and principal-agent relationships (in which managers run firms for the benefit of owners and creditors), a borrower's net worth becomes an important determinant of the amount of credit available, even though it is the economic prospects of the project being financed that should determine the availability of financing. When the net worth of a firm deteriorates, its supply of credit may be curtailed quite sharply, even when it has viable new investments. Again,

the generally high leverage of East Asian corporations could have exacerbated the impact of a credit crunch. Because interest payments account for a large share of the cash flow of such firms, small shocks would have had a large effect on borrowers' net worth.

Several papers have tried to model these relationships in a domestic context (Agénor and Aizenman 1999; Chan-Lau and Chen 1998; Kim and Stone 1999) and in an international context (Greenwald 1999). Generally, weaknesses in the financial sector along with tighter regulation and supervision appear to have contributed to corporate distress. There has been a presumption, with some supporting empirical evidence, that the tighter rules for financial institutions affected the supply of loanable funds in several countries (Ding, Domaç, and Ferri 1998; Ghosh and Ghosh 1999). This credit contraction, in turn, was likely to have led to a decline in output. In some countries and over time, the importance of credit contraction may have been diminished by the reduced demand for funds as the financial situation deteriorated and the prospects for returns on new investment worsened. Indeed, Ghosh and Ghosh (1999) find that in Korea and Thailand (but not in Indonesia), the binding constraint was the slowing demand for funds rather than an inadequate supply of credit. Although survey results confirm the importance of access to credit (see Dollar and Hallward-Driemeier in this issue), it is not clear to what extent the lack of access to credit reflected the poor prospects of the firms and the state of the financial sector.[2]

Although all corporations were affected to some degree, those with riskier financing patterns were more likely to be denied credit. Dollar and Hallward-Driemeier provide some support for this argument, showing that smaller firms that did not rely as much on foreign exchange financing faced a less severe credit crunch. But the degree of informational asymmetries, as reflected in the type of financing firms obtained, probably mattered for continued access to external financing. Domaç and Ferri (1999) find that in Korea, small businesses were particularly vulnerable to shocks that affected smaller regional banks, which normally provided them with most of their funding.

Changes in net worth were also likely to have been important in reducing the supply of financing. East Asian corporations had relatively high levels of debt before the crisis and were thus at risk of easily depleting their net worth. Even relatively small shocks could have lowered net worth significantly and thereby led to a sharp cut in external financing. Harvey and Roper (1999), who analyze the effects of risky financial structures on subsequent East Asian corporate performance, argue that corporate managers bet their companies by trying to offset declining profitability with ever-increasing amounts of borrowing in foreign currencies. Those bets turned sour when local currencies plunged and companies could not generate enough funds to pay their debts. Claessens, Djankov, and Ferri (1999) estimate that about 30 percent of corporations in crisis-affected East Asian countries were insolvent in the fall of 1998. More than twice as many corporations in these countries suffered illiquidity.

With banks mired in bad debts and external financing scarce, firms had to cut back on production and sales, including exports.

Even operationally viable firms were overburdened with debt, and banks were unwilling to provide capital to these firms until their debts had been resolved. This process can take a considerable amount of time where there are multiple creditors and weak frameworks for restructuring failed entities. Meanwhile, corporate value may be lost. The rapid rise in the share of nonperforming loans in bank portfolios also suggests the importance of financial shocks in precipitating the crisis, although the numbers do not allow one to differentiate the causes of the nonperforming loans (aggregate versus idiosyncratic shocks). More generally, a systemic financial crisis can exacerbate the effects of market imperfections (Greenwald 1999; Stiglitz 1999) As a result, firms that are operationally viable but financially distressed may suffer from a lack of working capital and other financing and be unable to maintain ongoing operations.

These explanations do not contradict the fact that many firms confronted problems arising from high leverage and low (risk-adjusted) rates of return and were thus very much at risk. One explanation of the preference for high levels of debt—apart from the need for external financing to maintain high growth rates—may be that it enabled large shareholders to retain control over firms' operations and thus to continue to benefit from a disproportional share of firms' cash flows. Had firms instead relied on equity financing, the bondholders would have controlled the decision to expand into new businesses, based on clear, objective criteria. Claessens, Djankov, and Lang (2000; forthcoming b) find that many firms in East Asia were controlled by a few large shareholders. Two-thirds of those controlling owners also held senior management positions and thus had ample opportunity to divert profits and indulge in high levels of risk. In cases where only a few owners held all or substantially all control, greed was a strong incentive. Empirical evidence reveals that minority rights were often violated in firms controlled by inside shareholders. Valuations of such companies were far below those of comparable firms, suggesting large-scale expropriation (Claessens and others 1999b).

Ownership structures may also have mitigated or exacerbated the impact of the shocks. Claessens and others (1999a) show that ownership structures can explain excessive diversification, which lowered firms' market valuation before the crisis. Group affiliation may also have been a factor because it could have allowed for more diversification of risk. Indeed, Claessens and others (1999a) find that the market valuation of companies associated with groups was 3 percent higher after the crisis than that of nonaffiliated firms, suggesting that there were some benefits from diversification within the group. It appears that group affiliation gives rise to tradeoffs: lower performance in normal times, but some risk diversification in turbulent periods. At the same time, group affiliation was found to be associated with expropriation.

Hypothesis 4

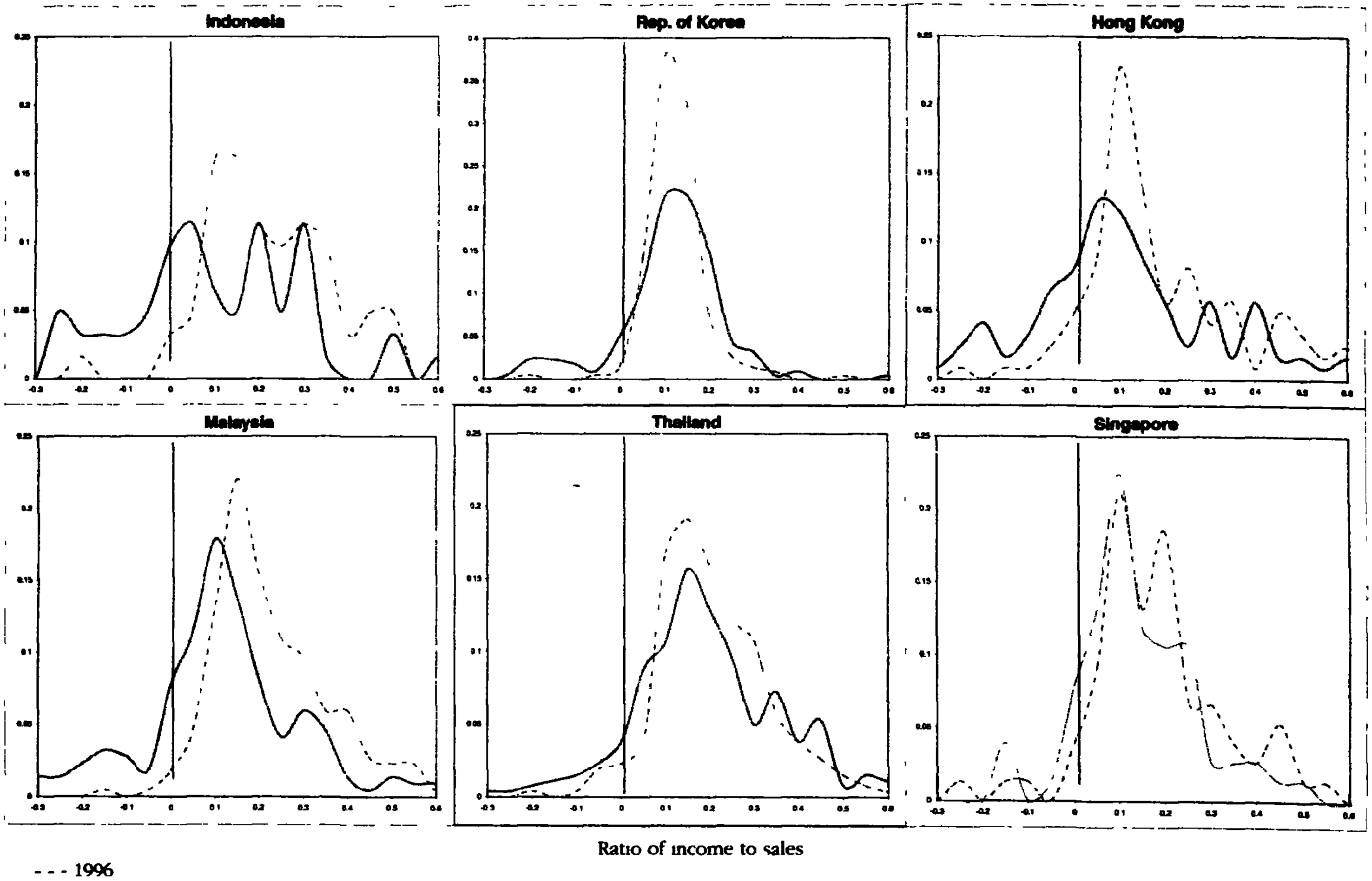
The final hypothesis states that the real effects of financial and other shocks depend in part on the efficiency of debt resolution mechanisms. Hart (1999) highlights the importance of the institutional framework in preventing and resolving systemic financial distress and discusses the possible need for extraordinary mechanisms during periods of systemic crisis. Stiglitz (1999) specifically argues that the optimal mechanism in such circumstances may be a bankruptcy or reorganization system that temporarily favors debtors and allows them to reorganize their operations. This literature also includes studies on the importance of creditors' rights to enforce claims and recover collateral as a means of facilitating financing in times of distress. More general work has focused on the importance of creditors' rights in determining whether firms have access to external financing (see La Porta and others 1999 for a review).

Frameworks for resolving financial distress differed considerably across countries, and these differences could have influenced the ability of firms to maintain operations and profitability. Differences in the institutional framework also appear to have influenced the actual use of bankruptcy procedures. Except for Korea and Malaysia, large corporations in the East Asian economies have made little use of formal bankruptcy procedures; more use has been made by all firms of the out-of-court systems that countries adopted after the crisis began. Indeed, Claessens, Djankov, and Klapper (1999) find that the strength of creditors' rights and the enforcement of these rights influenced firms' decisions to file for bankruptcy as a means of resolving financial distress.

The Performance of Publicly Traded Corporations

To provide some quantitative insight into the impact of the various factors in the financial crisis, we analyzed the performance of a subset of corporations before and after the crises in Indonesia, Korea, Malaysia, and Thailand and in two comparator countries, Hong Kong and Singapore. The data are from the Worldscope database, but in this analysis, the data set is balanced (that is, the same number of observations is used in 1996, 1997, and 1998). Because we include data for 1998 and need a balanced data set, we have to rely on a smaller set of firms for each country. As a measure of performance, we use firms' profit margin on sales, calculated as the earnings before interest and taxes plus depreciation and amortization, divided by total sales. The advantage of this calculation is that, as a cash-flow measure, it excludes interest payments, financial incomes, and other income and expenses and is therefore not influenced by the liability structure of the corpora-

Figure 3. *Ratio of Operating Income to Sales, Selected Economies, 1996 and 1998*



tion. This makes it possible to study the effects of real and financial shocks on operational performance.

The data set consists of 857 firms, of which 104 are in Hong Kong, 50 in Indonesia, 219 in Korea, 191 in Malaysia, 71 in Singapore, and 222 in Thailand. The data cover firms of different sizes; the median size is 1,099 employees, with the largest company employing almost 40,000 people and the smallest only 38. Sales volumes range from slightly less than $1 million in 1996 U.S. dollars to about $6.6 billion; the mean sales volume is $68 million. Overall, the data set covers primarily large firms, mainly because they have to be listed on a stock exchange to be included in the database and publicly traded companies tend to be large.

The postcrisis deterioration in performance across countries and firms can be illustrated by plotting the margins by country for all firms both before and after the crisis (figure 3). Note that the distributions shift to the left as margins deteriorate for all countries and that the number of firms with negative margins increases sharply in all countries. For the median firm, the margin on sales falls by 6.7 percentage points. The most seriously affected firms were in Indonesia, where median margins fell by 11.4 percentage points. The variability of margins across firms in each country also increased significantly, as shown by the distributions in 1998, which are wider (more fat-tailed) than the bell-shaped curves of 1996. The larger variability highlights the importance of analyzing the contribution of various groups of factors to individual firm performance.

To illustrate the importance of the various factors in explaining changes in performance, we run regressions using the firms' margin in 1998 as the dependent variable. The control variables are divided into four groups, corresponding to the roles of aggregate shocks, the nonfinancial characteristics of firms, the financial characteristics of firms, and country characteristics. The first group consists of industry and country dummy variables and aims to capture the country- and industry-specific shocks to which a firm may be exposed.[3] The dummies also capture some of the differences in market structures across industries, differences in foreign exchange depreciations, and other changes in the degree of relative competitiveness.[4] The second group includes the firms' nonfinancial characteristics before the crisis, including sales margin, real sales growth, and logarithm of sales (in U.S. dollars as a measure of size), all measured in 1996. The third group captures the firms' financial structure and ownership variables before the crisis, including the leverage ratio (defined as total debt divided by the market value of equity), the share of short-term debt in total debt, and ownership concentration (as measured by the percentage of control rights held by the two largest shareholders). The fourth group depicts the institutional environment of the country in which the firm operates. It includes indexes of the protection of shareholders' rights and creditors' rights and of the country's judicial efficiency in enforcing these rights.

We include these independent variables in a cross-sectional, pooled regression to try to explain the 1998 margins on sales. Specifically we estimate the following regression:

$$\begin{aligned} \text{margin}_{i,98} = {} & \alpha_0 + \Sigma_c \alpha_c D_c + \Sigma_j \alpha_j D_j \\ & + \beta_1 \text{ size} + \beta_2 \text{ margin}_{i,96} + \beta_3 \text{sales growth}_{i,96} \\ & + \gamma_1 \text{ ownership concentration} + \gamma_2 \text{leverage}_{i,96} + \gamma_3 \text{ short-term debt ratio}_{i,96} \\ & + \delta_1 \text{ judicial efficiency index}_c \times \text{equity rights index}_c \\ & + \delta_2 \text{ judicial efficiency index}_c \times \text{creditors' rights index} + \varepsilon_{it} \end{aligned}$$

where i indicates firm I, j refers to the industry, and c is a country index. Greek letters indicate coefficients to be estimated. To facilitate interpretation, we also standardize the firm- and country-specific variables to obtain a normal distribution with a mean of 0 and a variance of 1.5.[5] The results for different specifications are contained in table 5.

The first specification (column 1) controls for country and industry characteristics. Differences in 1998 margins across countries and industries are reflected in the coefficients. For countries, and controlling for other factors, the regression shows that firms in Thailand had significantly higher margins, 1.5 percentage points higher, than did firms in Hong Kong, the base economy. Among industry classifications, the commerce and construction industries had statistically significant lower margins compared with manufacturing, the base category, while the service industry had statistically significant higher margins. These cross-industry differentials in margins suggest that some of the causes of firms' financial distress were sector-specific shocks.

The next specification (column 2) includes firm characteristics before the crisis to help explain postcrisis margins. We examine both nonfinancial and financial characteristics of firms. Not surprisingly, firms with higher margins and sales growth before the crisis also had higher margins afterward; that is, firms that were performing relatively better before the crisis were also less affected by the crisis. Put differently, underlying prior weaknesses may have been a factor in firms' weak performance in 1998. This finding extends even to larger firms, which might have been expected to fare better than smaller ones because of greater diversity of products or preferential access to financing. However, postcrisis performance is not consistently affected by firm size (measured by the logarithm of total sales in dollars in 1996), suggesting that larger firms were not necessarily better able than smaller ones to weather the crisis.

In this specification, ownership concentration, as measured by the share held by the top two owners, has no significant relationship to postcrisis performance. Although insider control and connected lending have often been blamed for the problems of East Asian corporations, firms with concentrated ownership structures appeared to be no less able than other firms to manage the crisis. Other financial

characteristics did play a role, however. In particular, firms with higher leverage and a higher proportion of short-term debt tended to perform more poorly (and the coefficients are statistically significant in the median regression) than did firms without those characteristics. This finding is consistent with the view that the financial structures of East Asian firms before the crisis contributed to their poor performance afterward, suggesting that financial market imperfections and the credit crunch were important factors leading to lower operational performance.

The last specification (column 3) relates the postcrisis performance of a company to its country's institutional environment. These regressions drop the country dummies to focus on institutional differences only. Most of the coefficients found to be statistically significant in earlier regression results retain their significance here. We

Table 5. *Empirical Results*
(dependent variable 1998 profit margin divided by sales)

	(1)	*(2)*	*(3)*
Constant	0 048	0 089	0 134
	(0 016)**	(0 010)**	(0 006)**
Country			
Singapore	0 039	0 040	
	(0 024)	(0 015)**	
Malaysia	0 026	0 007	
	(0 019)	(0 013)	
Korea, Rep of	0 049	0 059	
	(0 018)**	(0 014)**	
Indonesia	–0 007	–0 033	
	(0 026)	(0 018)	
Thailand	0 118	0 085	
	(0 018)**	(0 015)**	
Sector			
Diversified	0 014	–0 009	–0.018
	(0 024)	(0 016)	(0 019)
Commerce	–0 057	–0.059	–0.071
	(0 017)**	(0 011)**	(0 013)**
Utilities	0 027	–0 040	–0 030
	(0 021)	(0.014)**	(0 017)
Services	0 127	0 002	0 011
	(0 022)**	(0.015)	(0.017)
Agricultural goods	–0 043	–0 004	0 003
	(0.025)	(0 017)	(0 020)
Mining	0 023	0 028	0 030
	(0 054)	(0.035)	(0 042)
Construction	–0 048	–0.052	–0.043
	(0 022)*	(0.014)**	(0 017)*

(*Table continues on the following page.*)

Table 5 *(continued)*

	(1)	*(2)*	*(3)*
Organizational factors			
A. Log (sales 96)		0.003	0.002
		(0.004)	(0.005)
Margin 96		0 099	0.098
		(0.004)**	(0.004)**
Sales growth 96		0.009	0.008
		(0.002)**	(0.002)**
Financial structure			
B Ownership concentration (top two owners)		0 005	0.002
		(0.006)	(0.005)
Leverage 96		–0.015	–0.014
		(0.003)**	(0 003)**
Short-term debt 96		–0.011	–0.013
		(0 004)**	(0.004)**
Institutional environment			
C Equity rights × judicial efficiency			0.004
			(0.014)
Creditors' rights × judicial efficiency			–0.029
			(0 014)*
Observations	857	857	857
R-squared	0 04	0 13	0 12

* Significant at 5 percent

** Significant at 1 percent

Note The *R* squares reported for the median regressions are pseudo *R* squares

All firm- and country-specific variables (except the industry dummies) are standardized to variables with means of 0 and variance of 1

A constant (not reported) is included in all regressions

Definitions of variables labor is the total number of employees, sales is sales revenues in US$1,000, ownership concentration is the sum of ownership stakes of the largest two shareholders, margin 96 is EBITDA (earnings before interest and taxes, depreciation added) sales in 1996, sales growth 96 is the real sales growth rate in 1996, leverage 96 is the sum of short-term and long-term debt over the market value of equity in 1996, diversified is a dummy variable indicating whether the firm operates in more than one 2-digit SIC industry, creditors' rights is an index measuring the protection of creditors, which ranges from 0 to 4 and is taken from La Porta and others (1999), shareholders should have "equity" rights, judicial efficiency is an index of the quality of judicial enforcement, which ranges from 0 to 10 and is taken from La Porta and others (1999)

Source Worldscope database

nteract the judicial efficiency index with indexes of equity rights and creditor rights o account for the combined effect of the strength of the laws and the quality of their enforcement. The results suggest that firms in countries with better protection of equity rights had better postcrisis performance. An increase in the equity rights index s associated with higher margins, although the coefficient is not statistically significant at conventional levels. This suggests that in countries with better equity rights, irms might have been better disciplined by owners and thus might have had structures and operations better able to withstand the shock of the crisis. A higher level of he effective creditors' rights index in the country is associated with poorer firm performance. The negative and significant coefficient for the interaction term between creditors' rights and the judicial efficiency index suggests that a stronger protection of creditor rights coupled with an efficient judicial system contributed to lower postcrisis margins. This finding may be consistent with the notion that when irms' financial distress is caused by exogenous external shocks, an overly vigorous protection of creditor rights might result in too many bankruptcies and losses in output.

We analyzed the contribution of each set of variables in explaining the percentage of variations in firm performance after the crisis. Firm-specific nonfinancial characteristics are the most important in explaining the variations in margins in 1998, accounting for roughly 5.7 to 6.3 percent of total variation, or about one-half to two-thirds of all variation that was explained. A firm's financial structure and industry affiliation were roughly equally important, each accounting for between 1.1 and 3 percentage points of total variation. The least important factor was the country's institutional environment, which contributed between 0.5 and 1.6 percentage points of total variations in margins in 1998.

In sum, the regressions were unable to explain nearly nine-tenths of the total variations, suggesting that the effects of idiosyncratic shocks, rather than aggregate shocks, were very large. Of the variance that could be explained, the results suggest that firm-specific characteristics, both financial and nonfinancial, were the most important in explaining postcrisis performance.

Although the financial fragility of the corporate sector may not have triggered the crisis, it did contribute to its depth and severity. This suggests that countries and the international financial community need to pay more attention to the status of the corporate sector and its links with the financial sector and the rest of the economy. One option would be to use balance sheet and other financial information to monitor financial risks in the corporate sector, but such surveillance will be limited by the paucity of data, the potentially rapid changes in corporate financial structures, and the analytical difficulties in identifying risk factors and linkages. The East Asian financial crisis showed that risks arising from the corporate sector typically occurred because of institutional weaknesses, including weak property rights, poor bankruptcy and accounting procedures, lack of transparency, and weak or perverse incentives.

Corporate sector vulnerabilities thus might be more successfully limited by assessing and remedying these deficiencies and the resulting weak risk management practices

Notes

At the World Bank, Stijn Claessens is lead economist in the Financial Sector Policy Group, Simeon Djankov is a financial economist in the Financial Sector Policy Group, and Lixin Colin Xu is an economist in the Development Research Group The authors would like to thank Ying Lin for able research assistance and the reviewers for helpful comments.

1 We present the share of long-term, rather than short-term, debt because the latter can underestimate the amount of liabilities with a short maturity, short-term debt excludes, for example, trade credits.

2. Krugman (1999) argues that the corporate balance sheets may have been at fault, apart from macroeconomic or other weaknesses. In particular, a depreciation of the domestic currency causes an increase in the currency value of foreign-denominated firm debt, at the same time that firms face declining sales and higher interest rates. The resulting balance-sheet problems and reversal of capital flows weaken the corporate sector and, in turn, the financial system This triggers a further currency depreciation with a current account surplus to accommodate the capital reversal deficit and financial system weakness Krugman ascertains that the risks of such an event occurring are higher when corporate profitability is low relative to the cost of funds to financial institutions

3 We classify firms into the following industry groups. commerce, utilities (including communications, electric, gas, transport, and water and sanitation services), services, mining, construction, agriculture and manufacturing When a firm is in more than one industry, we classify it as diversified

4 They do not, however, correct for differences in market structures and (potential) competition across countries. To the extent that firms produce traded goods, these country differences should not be important, but they will be important for other goods Also, the regressions do not try to control for the effects of exchange rate changes, which differed by country, on firm's performance measures

5 Because the margin and some right-hand-side variables have outliers, we conduct median regressions, which provide more robust estimates than ordinary least squares

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems

Agénor, Pierre-Richard, and Joshua Aizenman. 1999. "Financial Sector Inefficiencies and Coordination Failures Implications for Crisis Management " Policy Research Working Paper 2185 World Bank, Economic Policy and Poverty Reduction, World Bank Institute, Washington, D C

Alba, Pedro, Stijn Claessens, and Simeon Djankov 1998 "Thailand's Corporate Financing and Governance Structures. Impact on Firms' Competitiveness." In Johanna Witte and Stefan Koeberle, eds., *Competitiveness and Sustainable Economic Recovery in Thailand* Washington, D.C.. World Bank.

Bernanke, Ben, and Mark Gertler 1995 "Inside the Black Box. The Credit Channel of Monetary Policy Transmission " *Journal of Economic Perspectives* 9(4).27–48

BIS (Bank for International Settlements). 1999. "Capital Requirements and Bank Behaviour The Impact of the Basle Accord " Working Paper 1 Basle Committee on Banking Supervision At *http.//www.bis.org/publ/index htm*.

Chan-Lau, Jorge A., and Zhaohui Chen. 1998 "Financial Crisis and Credit Crunch as a Result of Inefficient Financial Intermediation—with Reference to the Asian Financial Crisis " IMF Working Paper WP/98/127 International Monetary Fund, Washington, D.C. Processed.

Cho, Yoon Je, and Joon-Kyung Kim. 1995. *Credit Policies and the Industrialization of Korea.* World Bank Discussion Paper 286. Washington, D.C.

Claessens, Stijn, and Thomas Glaessner. 1997 *Are Financial Sector Weaknesses Undermining the East Asian Miracle?* Directions in Development series Washington, D C.. World Bank.

Claessens, Stijn, Simeon Djankov, and Giovanni Ferri, 1999 "Corporate Distress in East Asia. Assessing the Impact of Interest and Exchange Rates Shocks." *Emerging Markets Quarterly* 3(summer):4–9.

Claessens, Stijn, Simeon Djankov, and Leora Klapper 1999 "Resolution of Corporate Distress: Evidence from East Asia's Financial Crisis " Policy Research Working Paper 2133. World Bank, Financial Sector Practice Department, Washington, D.C

Claessens, Stijn, Simeon Djankov, and Larry H P. Lang. 2000 *East Asian Corporations. Heroes or Villains?* World Bank Discussion Paper 409 Washington, D C

———. Forthcoming a. "East Asian Corporations: Growth, Financing and Risks over the Last Decade " *Emerging Markets Quarterly.*

——— Forthcoming b "The Separation of Ownership and Control in East Asian Corporations " *Journal of Financial Economics.*

Claessens, Stijn, Simeon Djankov, Joseph P H Fan, and Larry H. P Lang. 1999a. "Corporate Diversification in East Asia· The Role of Ultimate Ownership Structure and Group Affiliation " Policy Research Working Paper 2089. World Bank, Financial Operations Vice Presidency, Washington, D.C

———. 1999b "Expropriation of Minority Shareholders Evidence from East Asia." Policy Research Working Paper 2088. World Bank, Financial Operations Vice Presidency, Washington, D C

Colaço, Francis X , Mary Hallward-Driemeier, and Dominique Dwor-Frecaut 1999 "Asian Corporate Recovery A Firm-Level Analysis " World Bank, Washington, D.C. Processed

Corsetti, Giancarlo, Paolo Pesenti, and Nouriel Roubini. 1998 "What Caused the Asian Currency and Financial Crises? A Macroeconomic Overview " NBER Working Paper 6833. National Bureau of Economic Research, Cambridge, Mass

Demirguc-Kunt, Asli, and Vojislaw Maksimovic 1998. "Law, Finance, and Firm Growth." *Journal of Finance* 53(6) 2107–37

Ding, Wei, Ilker Domaç, and Giovanni Ferri. 1998. "Is There a Credit Crunch in East Asia?" Policy Research Working Paper 1959 World Bank, East Asia and Pacific Region, Washington, D.C.

Domaç, Ilker, and Giovanni Ferri 1999 "The Credit Crunch in East Asia Evidence from Field Findings on Bank Behavior." World Bank, East Asia Finance Group, Washington, D.C. Processed.

Furman, Jason, and Joseph Stiglitz 1998 "Economic Crises· Evidence and Insights from East Asia " *Brookings Papers on Economic Activity* 2.1–135

Ghosh, Swati R., and Atish R Ghosh. 1999 "East Asia in the Aftermath Was There a Crunch?" IMF Working Paper WP/99/38. International Monetary Fund, Washington, D.C. Processed

Greenspan, Alan, 1999. "Lessons from the Global Crises." Speech at the World Bank–IMF Annual Meetings, Washington, D C. At *http://www bog frb fed.us/BoardDocs/speeches/1999/199909272.htm*

Greenwald, Bruce 1999 "International Adjustment in the Face of Imperfect Financial Markets " In *Proceedings of the World Bank Annual Conference on Development Economics 1998.* Washington, D.C World Bank

Hart, Oliver. 1999 "Different Approaches to Bankruptcy " Massachusetts Institute of Technology, Economics Department, Cambridge, Mass. Processed

Harvey, Campbell, and Andrew H. Roper. 1999. "The Asian Bet." In Alison Harwood, Robert E. Litan, and Michael Pomerleano, eds., *Financial Markets and Development: Preventing Crises in Emerging Markets.* Washington, D.C.: Brookings Institution Press.

Johnson, Simon, Peter Boone, Alasdair Breach, and Eric Friedman. Forthcoming. "Corporate Governance in the Asian Financial Crisis, 1997–8." *Journal of Financial Economics.*

Kaminsky, Graciela, and Carmen Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." *American Economic Review* 89(3). 473–500.

Kashyap, Anil, and Jeremy Stein. 1994 "Monetary Policy and Bank Lending." In G. Mankiw, ed. *Monetary Policy.* Chicago, Ill.: University of Chicago Press.

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobatón. 1999 "Governance Matters." Policy Research Working Paper 2196 World Bank, Macroeconomics and Growth, Development Research Group, and Governance, Regulation, and Finance, World Bank Institute, Washington, D.C.

Kim, Joon-Kyung. 1999 "Creditors and Chaebols Relations in Korea" Korea Development Institute, Seoul Processed

Kim, Se-Jik, and Mark R. Stone. 1999. "Corporate Leverage, Bankruptcy, and Output Adjustment in Post-Crisis East Asia." IMF Working Paper WP/99/143 International Monetary Fund, Washington, D.C

Krugman, Paul. 1999. "Balance Sheets, the Transfer Problem, and Financial Crises" Massachusetts Institute of Technology, Cambridge, Mass. At *http://web.mit edu/krugman/www/#other*

Laeven, Luc 1999. "Risk and Efficiency in East Asian Banks." Policy Research Working Paper 225' World Bank, Financial Sector Strategy and Policy Department, Washington, D.C

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer 1998. "Law and Finance" *Journal of Political Economy* 106:1113–55.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny 1999 "Investor Protection. Origins, Consequences, Reform" NBER Working Paper 7428:1-42 National Bureau of Economic Research, Cambridge, Mass

Radelet, Steven, and Jeffrey Sachs 1998 "The East Asian Financial Crisis. Diagnosis, Remedies, Prospects" *Brookings Papers on Economic Activity* 1:1–90

Stiglitz, Joseph E 1999 "Bankruptcy Laws Some Elementary Economic Principles" Paper presented at the Annual Bank Conference on Development Economics, Europe 1999, Paris Processed

Stiglitz, Joseph E., and Amar Bhattacharya 2000 "Underpinnings for a Stable and Equitable Global Financial System. From Old Debates to a New Paradigm." *Proceedings of the World Bank Annual Conference on Development Economics 1999* Washington, D.C: World Bank.

Stiglitz, Joseph E, and Andrew Weiss 1981 "Credit Rationing in Markets with Imperfect Information" *American Economic Review* 71(3): 393–410.

Industrial Policy: Growth Elixir or Poison?

Howard Pack

The debate about industrial policy occasioned by the East Asian financial crisis is the latest chapter in an ongoing discussion about the effectiveness of selective government intervention in fostering rapid industrial growth. The crisis that began in the Republic of Korea in 1997 and the weak growth in Japan over much of the 1990s have prompted a reexamination of the effectiveness of the government actions in the two countries that pursued sectoral selectivity most intensively. If indeed industrial policies were important in accelerating growth, there may be lessons for other countries still in the early stages of industrialization. Conversely, if the magnitude of the contribution was small, more conventional policies should be pursued unless it is assumed that governments can improve on the efforts of Japan and Korea.

The East Asian financial crisis that erupted in July 1997 renewed interest in the East Asian miracle, a term that had only recently entered the development vocabulary. Many causes were identified quickly, ranging from the excessively rapid opening of capital markets to the decline in the extent of prudential regulation of the banking system that had characterized the countries through the early 1990s. Some analysts argued that the financial vulnerabilities revealed during the crisis stemmed in part from earlier efforts to implement an industrial policy that required commercial banks to lend to sectors and firms selected by the government. Banks, in this view, had inherited weak balance sheets as a consequence of the low profitability of these preferred sectors (see Claessens, Djankov, and Lang 1998). Moreover, the banks had lost (or failed to acquire) the ability to select, evaluate, and monitor individual loans (Fry 1995). The huge capital inflows that began in the mid-1990s into countries such as the Republic of Korea thus were channeled through institutions that had suffered a serious erosion of skills and discipline. Industrial policy, in this view, had been a slow-acting poison for the countries pursuing it, one that necessitated emergency measures in 1997 and 1998. Others argued that, on the contrary, industrial policy had been a growth elixir that led to more rapid growth than would have been

The World Bank Research Observer, vol 15, no 1 (February 2000), pp 47–67

obtained from simply following good macroeconomic policies combined with supportive efforts such as encouraging the rapid growth of education (Amsden 1989; Lall 1997).

The debate about industrial policy occasioned by the financial crisis was only the latest installment in an ongoing debate about the efficacy of government intervention in fostering more rapid economic growth. Some of the earliest literature in development economics, based on a preliminary understanding of Soviet growth, argued that sectors such as heavy metals and machinery were particularly important and that these sectors generated externalities whose benefits were so desirable that they warranted government support. Other analysts argued that there were important market failures such as real external economies stemming from research and development and important imperfections in the market for information (Chenery 1959; Scitovsky 1954). Partly in response to such views and partly because of the perception that agriculture was inferior to industry, many governments initiated import substitution policies in the 1950s.

By the 1970s import-substituting industrialization was widely viewed as a failure. Several rounds of empirical research showed that it led to static misallocation of labor and capital across industries and did not improve long-run growth in total factor productivity (Balassa and associates 1982; Bhagwati 1978; Krueger 1978; Little, Scitovsky, and Scott 1970). These interpretations were based on careful empirical research in a large number of countries, from Argentina to India. The rapid growth for three decades or more of a number of East Asian economies, including Japan, Korea, Singapore, and Taiwan (China), all of which pursued government-led industrial policies, gave rise to an optimism that such policies, if executed correctly, could have beneficial effects (Pack and Westphal 1986; Rodrik 1995; Stiglitz 1996). But the crisis that began in Korea in October 1997 and the weak growth in Japan over much of the 1990s have prompted a reexamination of the impact of industrial policy in those two countries. If indeed industrial policies were important in this process, there may be lessons for other countries still at early stages of industrialization. Conversely, if the magnitude of the contribution was small, more conventional policies should be pursued unless other countries assume that their governments can outperform those of Japan and Korea

Industrial policies comprise a variety of actions designed to target specific sectors to increase their productivity and their relative importance within the manufacturing sector. Proponents of an active industrial policy place considerable emphasis on the potential of such activity for encouraging a shift to newer and more modern sectors—sectors that are characterized by industrywide external economies and by learning-by-doing on the part of individual firms. Industrial policy is viewed as likely to lead to growth in total factor productivity (TFP) in the targeted sectors. Such arguments were used to rationalize import-substituting industrialization policies, but the evidence suggests that policymakers' hopes for success were not realized. The

question is whether the undoubtedly better policies pursued by Japan and Korea constitute a basis for optimism about the potential gains from a well-designed set of industrial policies.

In both countries, government actions that contributed to rapid economic growth extended well beyond industrial intervention: promoting education, building a large and efficient social infrastructure, encouraging international technology transfers, and fostering research and development. Macroeconomic policies that promoted growth included the encouragement of private saving, the maintenance of stable real exchange rates, the prevention of speculation in financial markets for a considerable part of the period (Stiglitz 1994), and a fiscal-monetary stance that limited inflation.

Interpretations of the Sources of Success

A market-oriented interpretation of the success of Japan and Korea had gained ascendance by about 1980 and contrasted the experience of the two East Asian countries favorably with that of Latin America and Southeast Asia (Little 1982). The core of this view was that after brief interludes of intensive import-substituting industrialization, Japan and Korea switched to increasingly liberal trade policies. Macroeconomic policy controlled inflation relative to the experience of many developing countries, and the real effective exchange rate rarely appreciated. When it did, such episodes were quickly corrected. Manufacturers were thus able to concentrate on improving productivity rather than coping with rapidly changing relative prices of inputs and outputs. In contrast, Latin American countries experienced considerable macroeconomic dislocations, from high rates of inflation to periodic balance of payments crises. These countries protected their industries for several decades and did not attempt to offset the adverse effects of such policies on exporters Moreover, for much of the period, exchange rates were volatile and overvalued. Thus attempts to foster specific industries occurred in a context that did not favor the learning necessary for manufacturers to become internationally competitive.

After a consensus had formed on the role of markets in fostering growth in Japan and Korea, additional research uncovered significant amounts of discriminatory policies in the two countries in both product and factor markets, although less than in other countries (for Korea, see Amsden 1989; Jones and Sakong 1980; Pack and Westphal 1986; for Japan, see Johnson 1982; Yamamura 1986; Yamamura and Yasuba 1985). For example, the extent of selective trade protection in individual sectors is now documented to have been quite substantial in Korea and Japan, although less than in other developing countries (for a survey on protection in various developing countries, see Erzan and others 1989). This has led to support for the view that selective government intervention, if well executed, could be a major contributor to growth. This article tests that hypothesis.

Japan targeted a number of industries in the 15 years after the end of World War II, in an effort to restore them to their previous levels of capacity and efficiency, and relied on tariff protection and subsidized interest rates to encourage the introduction of more advanced technology. At the same time, rationalization cartels facilitated the exit of higher-cost firms (Mutoh 1988; Tanaka 1988; Yamawaki 1988; Yamazawa 1988; Yonezawa 1988). These efforts to rebuild industries that were previously highly productive differed from the attempts that were launched in the 1960s to encourage new product areas. These later policies, which attempted to "pick winners," were pursued with a variety of policy instruments (Johnson 1982; Yamamura 1986). Evidence of varying levels of protection for Japan is shown in table 1.

In the 1960s Korea encouraged the development of industries such as textiles by protecting the domestic market and by offering export subsidies and access to low-cost credit to firms that reached export targets. These efforts were limited in extent and may have had some effect on growth (Westphal 1978). An entirely different order of magnitude characterized Korea's interventions in the 1970s, which encouraged the growth of basic metals, machinery, and chemical sectors (typically described as the heavy and chemical industry—HCI—drive). Table 2 shows that in 1978, at the height of the promotion effort, tariffs and quotas provided very high levels of protection to domestic manufacturers of consumer durables, electrical and nonelectrical

Table 1. *Effective Protection in Japan, 1963, 1968, and 1972*

Sector	*1963*	*1968*	*1972*
All manufacturing	32.3	24.2	14.4
Textiles	54.3	28.2	18.6
Clothing	72.8	32.8	22.4
Wood products	14.0	25.6	16.1
Pulp, paper	9.7	18.0	11.0
Publishing	-16.7	1.0	-0.9
Leather and rubber products	30.9	21.8	12.3
Chemicals	33.4	17.7	8.8
Petroleum and coal products	19.5	14.5	7.1
Nonmetallic minerals	22.2	15.7	8.1
Iron and steel	30.1	30.0	17.1
Nonferrous metals	30.4	34.1	22.1
Metal products	13.8	19.9	9.9
Machinery	36.7	20.0	7.7
General	23.0	14.5	8.7
Electrical	30.9	16.5	5.4
Transport equipment	61.5	31.0	9.2
Precision instruments	34.9	22.9	10.4

Note These are minimum estimates insofar as the calculations of effective rates of protection are based on tariff data, not on implicit price estimates of nominal levels of protection

Source Shouda (1982) as cited in Itoh and Kiyono (1988).

Table 2. *Effective Rates of Protection and Subsidy, Korea, 1978*

Sector	*Subsidy rate for exports*	*Effective rate of protection for domestic sales*	*Average incentive effect for domestic and foreign sales*
Processed food	31.7	–29.4	–23.0
Beverages, tobacco	13.2	28.0	27.8
Construction materials	19.1	–15.0	–10.5
Intermediate products 1	23.6	–37.9	–31.4
Intermediate products 2	26.3	7.9	12.0
Nondurable consumer goods	17.3	31.5	24.0
Consumer durables	38.0	131.2	83.2
Machinery	24.4	47.4	43.2
Transport equipment	26.1	135.4	87.2
All manufacturing	22.8	5.3	9.7

Note The two categories of intermediate products represent different levels of manufacturing complexity For precise industry definitions, see Nam (1990)

Source Nam (1990)

machinery, and transport equipment. Simultaneously, substantial loans were directed to the targeted sectors, with interest rate subsidies for the HCI sectors ranging from 2 to 6 percent (Sakong 1993: table A.19).

Were these attempts at promoting specific sectors essential to the success of the two East Asian countries, or were they a marginal gloss on the governments' nonselective macroeconomic policies? (See Stiglitz 1996 for an account of the nontraditional policy interventions and how they may have stimulated growth in Japan and Korea.) If the rapid growth of these countries was attributable to selective policies, other poor countries might learn from them. Skeptics argue, however, that Japan and Korea would have grown even faster had it not been for selective intervention. In this view, given macroeconomic stability, a realistic and stable real exchange rate, the availability of traded inputs at international prices, relatively competitive labor markets, very high saving and investment rates, high and growing levels of education, and considerable numbers of capable entrepreneurs, the contribution of intervention was negligible or harmful (Saxonhouse 1983; Trezise 1983; Wolf 1988).

Testing the Effects of Intervention

Economic analysts employ a variety of strategies to assess the quantitative impact of import substitution policies on individual economies. To establish the magnitude of static losses in output, researchers estimate effective rates of protection and derive a cost for misallocation of resources of 3–6 percent of gross domestic product (Pack 1988). Another approach calculates the growth rate of TFP during alternating periods of import restrictions and free trade (Krueger and Tuncer 1982). Using an elabora-

tion of the second approach, I calculate the combined impact of industrial policies on the sectoral structure of manufacturing and the TFP growth rate of individual sectors. If the calculated effect is large and positive, there may be a basis for least-developed countries, such as those in Africa, embarking on more-intensive industrialization, to weigh the benefits and costs of some elements of industrial policy. If, however, even Japan and Korea, where general policies were well designed and bureaucracies were very capable, were unable to extract major benefits from selective intervention to foster individual sectors, this would seem an unpromising path to industrialization.

To calculate the effects of industrial policy on the growth rate of total factor productivity for the entire manufacturing sector, I first calculate TFP growth in each sector weighted by the sector's share of value added:

$$\text{TFP}^* = \sum_i (VA_i/VA)\text{TFP}_i^* \qquad (1)$$

where i indicates individual branches such as food processing, textiles, and machinery within the manufacturing sector, VA_i is value added originating in sector i, VA is manufacturing-wide value added, and TFP_i^* is the rate of TFP growth in sector i. The ratio of value added in each sector to value added in all of manufacturing, VA_i/VA, represents the industrial structure. Industrial policy may affect manufacturing-wide productivity by increasing the importance of industries whose TFP growth is above average, by altering VA_i/VA, or by stimulating TFP_i^* itself. The major issue is to establish a counterfactual set of figures for the two variables contained in equation 1.

The ratios of value added in each sector to value added in all manufacturing can be thought of as representing industrial structure, where sectors with higher ratios are more prominent than sectors with lower ratios. TFP growth in all manufacturing, then, is the sum of TFP growth in each sector, weighted by a measure of the importance of that sector.

The Counterfactual Sectoral Structure of Production

I consider the periods 1960–79 for Japan and 1966–85 for Korea. The period 1960–79 is one in which Japan initiated efforts to enter new sectors. In Korea the period 1966–85 includes the intensive heavy and chemical industry program of 1973–80 Although the Korean government encouraged some sectors in all periods, the magnitude of the HCI effort dwarfed previous and subsequent efforts to promote individual sectors. I use the values of VA_i/VA at the end of the periods during which industrial policy was important to calculate a base value for equation 1. The use of end-of-period sectoral shares is favorable to finding a positive effect of industrial policy because it assumes that any higher growth rate of favored sectors that resulted from sectoral promotion generated benefits for the entire period rather than only part of it.

Without government intervention, international trade theory predicts that the major determinant of the structure of production (and international trade patterns) is the relative supply of factors of production, assuming that tastes are similar across countries. Macroeconomic and education policies may affect relative factor supplies, but these policies are not targeted to particular sectors and, by definition, are not part of industrial policy. Most of the research using these models has generated relatively poor predictions of production and trade patterns. Statistical estimates that use as many as 10 factors of production do not explain well the distribution of value added among manufacturing sectors (Harrigan 1995) or the pattern of exports (Noland 1997). Given the lack of a rigorous statistical model for generating the counterfactual values of VA_i/VA, I employ these values for a number of countries that had similar income per capita at the end of the periods in question and for which the value added by detailed industrial branches is available.

Japan's benchmark countries in 1979, in ascending order of per capita income, are Italy, Finland, Austria, Australia, France, and the Federal Republic of Germany (table 3). These countries are relatively similar to Japan in terms of their stock of capital and the education of their citizens, so that differences in industrial structure might be attributable to specific sectoral interventions. For Korea, the countries of comparison are Uruguay, Portugal, Malaysia, Mexico, and Panama. None of these had fostered the intensive industrialization of Korea; certainly none had attempted so consciously to shift the structure of production. Using these countries as peers provides a maximum measure of the alteration in the structure of production engendered by industrial policy. Other researchers have used alternative measures of sectoral evolution, including those implied by Syrquin and Chenery (1989). Use of the sectoral distribution implied by their estimates suggests a smaller impact on industrial structure in both Japan and Korea than in the peer countries. Thus the calculation used here contains a bias in favor of finding a greater impact from discrimination among industries.

To compare Japan and Korea with their peer countries, I calculate average sectoral shares of value added for each of 28 sectors in the peer economies and aggregate these to 11 sectors (figures 1 and 2). Japan and Korea show a few substantial sectoral differences in the structure of production relative to their benchmark countries. In particular, in both countries the combined metal products, machinery, and electrical equipment (MPMEE) sector is much larger than in the peer countries, suggesting that government policies may have had their intended impact on sectoral structure.

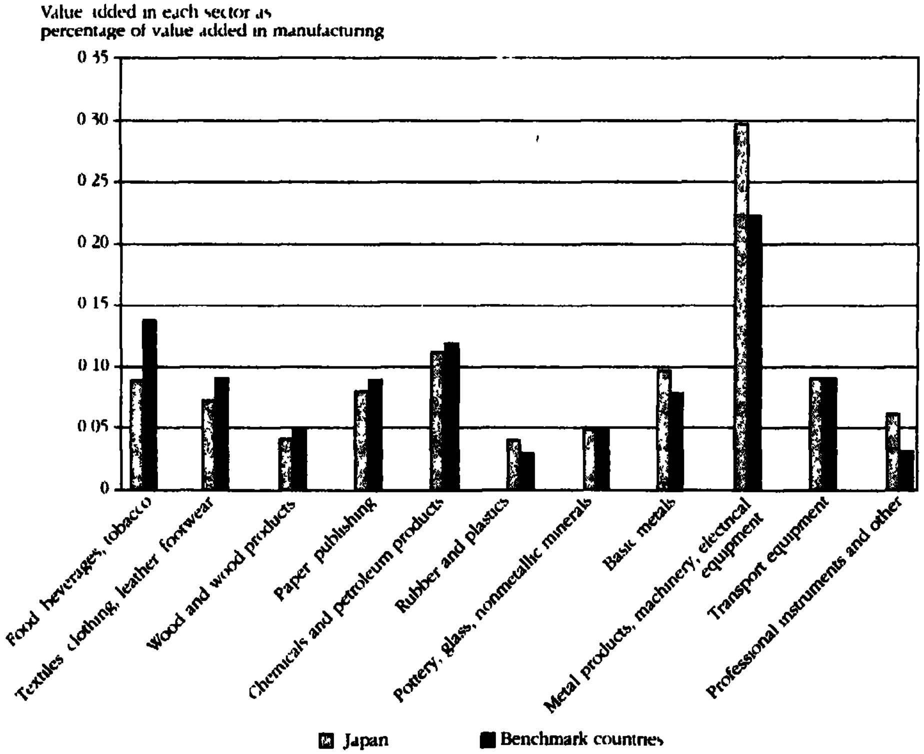
By using the industrial sectoral structure derived from the benchmark countries, I am assuming that all of the observed differences in industrial structure are attributable to industrial policy rather than to (unobserved) differences in supplies of factors of production. Insofar as the favored sectors in Japan and Korea contributed greater shares to total value added *and* exhibited higher rates of TFP growth, this assumption

Table 3. *Per Capita Income of Japan, Korea, and the Benchmark Countries, 1979 and 1985*
(U.S dollars in current prices)

Country and year	*Per capita income*
1979	
Japan	8,810
Italy	5,250
Finland	8,160
Austria	8,630
Australia	9,120
France	9,950
Germany, Fed. Rep.	11,730
1985	
Korea, Rep. of	2,150
Uruguay	1,650
Portugal	1,970
Malaysia	2,000
Mexico	2,080
Panama	2,100

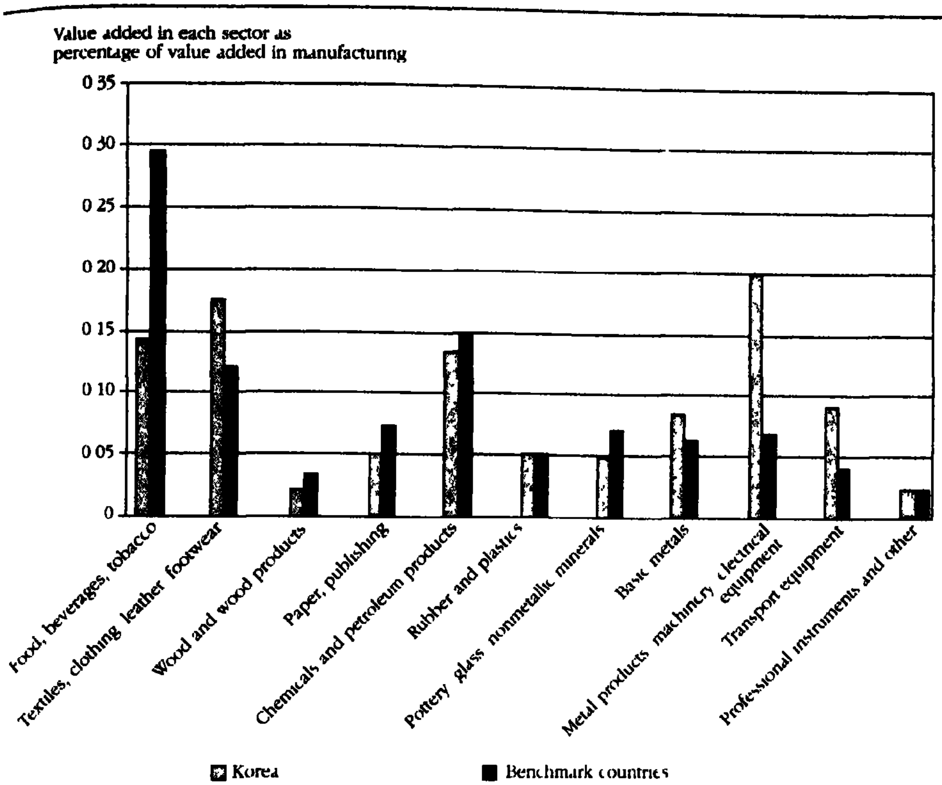
Source World Bank (various years).

Figure 1. *Sector Shares, Japan and Benchmark Countries*



Source United Nations (various issues)

'igure 2. ***Sector Shares, Korea and Benchmark Countries***



Source United Nations (various issues)

verstates the contribution of industrial policy. But the value calculated helps to ›tablish a plausible upper limit on the effect of industrial policy.

'he Counterfactual Impact of Industrial Policy on Manufacturing-Wide 'roductivity Growth

is possible to calculate the counterfactual growth of TFP by comparing actual TFP 'owth rates across sectors in Japan and Korea or by comparing TFP growth in each 'ctor in Japan and Korea with that in comparable countries. Differences in rates of roductivity growth across countries, however, reflect country-specific factors, such › the intensity of competition, public investment in infrastructure and education, ıe legal structure governing intellectual property rights, and firms' investments in arning and worker training. It is inappropriate to attribute the higher rates of TFP owth in favored sectors in Japan and Korea, compared with those in, say, Austria · Mexico, to selective industrial policy alone, since Austria and Mexico lack several

of these productivity-enhancing characteristics. I thus focus on interindustry differences in productivity in Japan and Korea by comparing favored and neglected sectors rather than relying on international comparisons for identical industrial branches.

Jorgenson, Kuroda, and Nishimizu (1987) have estimated sectoral TFP growth for Japan during 1960–79 (table 4).[1] As noted, many of the critical acts of the Japanese government with respect to industrial policy occurred in the 15 years following World War II and were designed to reestablish the previous peak levels of production in sectors such as coal mining, textiles, and steel production. The attempt to foster nontraditional industrial sectors began around 1960. It can be seen that MPMEE and transport equipment exhibited much higher rates of growth of productivity than other sectors, confirming the views of proponents of Japan's industrial policy. Given that the MPMEE sector also had a higher share of VA_i/VA in Japan than in the benchmark countries, this resulted in an increase in sectoral TFP growth. It appears that industrial policy in Japan may have increased both the importance and the productivity of the industry; the issue is the precise size of the increase.

For Korea, I calculate measures of TFP and convert them to annual growth rates (table 5).[2] During the height of the HCI program, 1970–78, the mean values of TFP growth in the HCI sectors, 4.9, substantially exceeded the group average of the "ne-

Table 4. *Total Factor Productivity Growth in Japan, by Industrial Sector, 1960–79*

Sector	*TFP growth rate (percent)*
Food	–1 20
Textiles	0 47
Clothing	1 98
Leather and shoes	1 03
Wood	2 81
Furniture	1 74
Paper	1.44
Printing	–0.18
Chemicals	3 36
Petroleum	–3.55
Rubber	1.02
Plastics	0.55
Pottery and glass	0.92
Iron and steel	2.86
Other metals	0 16
Metal products	3.41
Machinery	2.30
Electrical machinery	5.38
Transport equipment	4.32
Professional instruments	4.45
Other manufacturing	4.98

Source Jorgenson, Kuroda, and Nishimizu (1987)

glected" sectors, 3.1. In the succeeding period, the neglected sectors had a higher average value of TFP growth. Individual sectors in Korea—both neglected and favored—had high TFP growth rates. The difficulty of establishing the beneficial effects of promotion on productivity can be seen in one example: the much greater growth of TFP in the industrial chemical sector in the period 1966–70, before the HCI drive, than in the succeeding periods in which the sector was encouraged. The fact that its TFP growth rate for the entire 1966–85 period is the highest of any sector is attributable to its performance in the earliest period. The same is true of the transport equipment sector.

The results in table 5 are more favorable to the HCI strategy than are other estimates of growth. For example, Dollar and Sokoloff (1990), examining 1963–79, find that TFP growth rates were higher in labor-intensive industries than in capital-

Table 5. *Total Factor Productivity Growth in Korea, by Manufacturing Sector, 1966–85*

Industry	*1966–70*	*1970–78*	*1978–85*	*1966–85*
Light industries				
Food	6.5	–0.8	3.7	2.3
Beverages	10.3	4.3	2.1	4.0
Tobacco	13.8	2.0	5.2	4.5
Textiles	7.2	3.4	2.8	3.3
Clothing	9.9	3.8	4.5	4.2
Footwear, leather	6.2	4.4	–3.5	2.4
Wood	1.1	6.6	–1.1	2.5
Furniture	25.4	–2.4	0.7	4.7
Pulp, paper	7.4	4.0	0.3	2.7
Printing	–3.2	4.1	0.0	1.4
Other chemicals	4.2	6.8	–1.8	2.9
Petroleum, coal products	3.5	–0.1	2.1	1.3
Rubber products	3.1	6.8	1.0	3.5
Nonmetallic minerals	4.0	4.0	0.2	2.5
Other manufactures	13.2	–0.1	–2.1	2.5
Average	7.5	3.1	0.9	3.0
Heavy and chemical industries				
Industrial chemicals	22.7	9.0	1.3	6.6
Oil refining	13.3	–1.8	–7.5	0.1
Iron and steel	–0.4	3.8	4.0	2.5
Nonferrous metals	0.4	3.7	6.4	3.6
Fabricated metals	1.1	8.0	–3.6	2.8
Nonelectrical machinery	1.2	7.7	–1.7	3.0
Electrical machinery	5.3	5.7	0.4	3.3
Transport equipment	8.5	3.3	4.0	3.9
Average	6.5	4.9	0.4	3.2

Source Author's calculation of Tornqvist indexes based on data contained in Yoo (1990)

intensive ones, which include the HCI sectors.[3] They note that "whereas capital deepening does indeed seem to have been the dominant factor in the (growth of) highly capital-intensive industries, total factor productivity growth was the major contributor to gains in labor productivity in the rest of manufacturing" (Dollar and Sokoloff 1990: 310). Their results would lower the increase in sectorwide TFP growth due to selective policy.

What is most striking about manufacturing productivity in Korea is the high absolute values of TFP growth in most sectors. Although there are sectoral differences, the individual values and their average are extremely high by international standards (Nishimizu and Page 1987). They exceed the productivity growth achieved by the Organisation for Economic Co-operation and Development (OECD) countries in the same period. Although industrial policy achieved its goal of shifting factors among sectors, the cost of doing so was forgone high productivity growth in the neglected sectors. Unless the "neglected" sectors also somehow realized high productivity growth as a result of policies that implicitly discriminated against them, industrial policy cannot have been a major factor in Korean industrial growth. (The potential impact of spillovers from promoted to neglected sectors is considered below.)

It seems likely that government pressure to export was a major impetus for the relatively high rates of growth of TFP. In return for subsidized loans and various other incentives, the government set ever-increasing export quotas for each firm (Jones and Sakong 1980); the targeted quotas depended on previous export growth and assessments of future potential. Given Korean firms' high ratios of debt to equity, the credible threat of withdrawal of subsidized credit was a potent inducement to meet export targets. Assuming that in each period firms were equating marginal cost to world price, a 20 percent increase in the export target forced them to reduce their costs rapidly if they were to avoid a loss on the increment in exports. Another source of productivity growth was technical advice that exporters received from customers in the OECD countries (Rhee, Ross-Larson, and Pursell 1984).

The Counterfactual Calculations

I use four calculations to measure the potential effect of industrial policies in Japan and Korea.

BASE CALCULATION. The basis of comparison is the calculated rate of TFP growth employing equation 1, using the observed value added shares in 1979 (Japan) or 1985 (Korea) and the sectoral TFP growth rates shown in tables 4 and 5. This calculation assumes that the sectoral structure of production and any differences between TFP growth in favored and neglected sectors were attributable entirely to industrial policy. The values for TFP growth, 3.12 for Korea and 2.36 for Japan, are shown in row 1 of table 6.

I now consider three variations of the basic calculation of what the path of TFP growth would have been in the absence of industrial policy.

ALTERNATIVE 1. In the first calculation, I assume that the only effect of industrial policy was to shift the sectoral patterns of production; TFP growth in individual subsectors was not affected by the policy. I assume that the sectoral pattern in the absence of industrial policy would have been similar to that in the benchmark countries. Under these assumptions, Korea's TFP growth was 3.04 percent and Japan's was 2.04 percent. In other words, Korea obtained a 0.08 percentage point increment (3.12 – 3.04) and Japan obtained a 0.32 percentage point increase in annual TFP growth rates from the policy-induced difference in sectoral structure.

ALTERNATIVE 2. The second scenario assumes the reverse: that, instead of shifting the sectoral pattern of production, industrial policy only had an effect on the TFP growth rates of the sectors. I assume that half of the observed TFP growth rates in the favored sectors was attributable to industrial policy. Under these assumptions, the contribution of industrial policy to TFP growth was 0.72 percentage point in Korea and 0.89 percentage point in Japan.

ALTERNATIVE 3. The third calculation combines the assumptions of the first two. I assume that industrial policy affected both the sectoral structure and productivity growth in the individual sectors. These assumptions imply that industrial policy accelerated TFP growth by 0.60 percentage point in Korea and by 0.99 point in Japan.

These calculations overestimate the benefits from industrial policy because they assume that the entire sectoral deviation in value added from benchmark countries and any improvements in TFP growth are due entirely to intervention.

Were these (maximum) figures of decisive importance to aggregate economic growth in the two countries? Given that Korea's manufacturing sector accounted for about a third of gross domestic product (GDP) at the end of the period, industrial policies that contributed 0.60 point to the manufacturing growth rate would have accounted for no more than 0.20 point of aggregate growth (roughly 9 percent a year). For Japan, a similar calculation implies that the extra 0.99 percentage point in

Table 6. *Simulation of the Effect of Industrial Policy on Total Factor Productivity Growth*

Scenario	*Korea*	*Japan*
Base	3 12	2.36
Alternative 1	3.04	2.04
Alternative 2	2 40	1 47
Alternative 3	2.52	1.37

manufacturing would have contributed about 0.3 point to aggregate growth of 9 percent. Even if these figures were doubled, it would still be the case that industrial policy was a minor hormone rather than the magic elixir of aggregate growth.

There are many anecdotes relating to the accelerated development of individual companies. If, however, the growth of these companies has not been sufficient to affect the results at the level of disaggregation used here, they cannot have been, *in the aggregate*, quantitatively important sources of growth.

Externalities and Investment Coordination

So far, it has been assumed that selective industrial policies directly affected the targeted sectors and that high rates of productivity growth in the neglected sectors did not reflect spillovers from the promoted sectors. But industrial policies could have generated benefits in other sectors as a consequence of three developments: domestic production of intermediate goods with special characteristics that were not available internationally but improved productivity in the local purchasing firm; job transfers by a trained labor force from firms in targeted sectors to other firms, bringing with them uncodified knowledge; and direct interactions on equipment design by producers and local buyers that led to adaptations that were particularly suitable for local firms.

All three externalities could potentially increase TFP growth in the neglected sectors. The potential quantitative importance of specialized nontraded intermediate inputs and uncodified knowledge transmitted by workers depends on how much the neglected sector interacts with the promoted sector. One way to gauge these benefits is to measure the purchases of an input from a favored sector per dollar of gross output in the neglected sector. The larger the purchase, the more likely it is that the neglected sector will benefit from the existence of local producers. The neglected sector may also derive greater benefits if there are few imports, which constitute an alternative source of specialized inputs.

Several measures showing the magnitude of the interaction between promoted and neglected sectors in Korea are presented in table 7; those for Japan are presented in table 8. First, in both countries the average input-output interaction is quite small. The favored sectors account for a very small portion of the domestically purchased inputs of most neglected sectors. Second, the heavy industries purchase extensively from one another. Third, in Korea imports by the neglected sectors are, on average, twice as large as total purchases from the favored domestic sectors (0.134 versus 0.068). In Japan, imports constitute a smaller percentage of total purchases. Whether this is due to the nontraded characteristics of Japanese production or to the restrictive trade regime is not clear (Lawrence 1993; Saxonhouse 1993). These patterns suggest the following probable effects of industrial policy on neglected sectors:

Table 7. *Intersectoral Purchases, Korea, 1985*

Purchasing sector	*Purchases from* All domestic sectors	Heavy industries	Chemical industry	Foreign suppliers
Neglected sectors	0.293	0 021	0.047	0 134
Favored sectors				
Chemical industry	0.357	0 010	0 249	0.209
Heavy industries	0.411	0.342	0 021	0 176

Source Author's calculations from input-output tables contained in Bank of Korea (various issues)

- It is unlikely that the promoted sectors were quantitatively critical in increasing the range of available inputs. Although industrial policies may have encouraged the domestic production of some unique, nontraded inputs, the overall impact was slight relative to all domestic and foreign purchases. Unless there was very low substitutability between domestic and foreign inputs, the quantitative effect of domestic inputs was small Rosenberg (1976) cites the importance of local interactions in situations where both user and producer were themselves at the world frontier and there were no suppliers in other countries. In contrast, Korean and Japanese firms in the periods considered were not at the world frontier in the neglected sectors and had many opportunities for obtaining specialized inputs from abroad.
- Insofar as workers and managers transmitted important knowledge, the small purchases from the promoted sectors imply that the effect of such learning was limited. Although one can posit, as in the case of specialized inputs, that there was a critical piece of knowledge whose possession had exceptionally high marginal productivity for the recipient sector, the quantitative case does not seem plausible. Moreover, such information could have been obtained from technology licensing agreements and foreign consultants.
- Promoted sectors are substantial purchasers of one another's inputs, at least in the metal-based sectors. Any externalities from such interaction are already accounted for in the calculations shown in table 6 insofar as they employ the

Table 8. *Intersectoral Purchases, Japan, 1980*

Purchasing sector	*Purchases from* All domestic sectors	Heavy industries	Chemical industry	Foreign suppliers
Neglected sectors	0.354	0 016	0 048	0 113
Favored sectors				
Chemical industry	0.532	0.011	0.356	0.051
Heavy industries	0.521	0 329	0.008	0 035

Source Author's calculations from input-output tables contained in Bank of Japan (various issues)

Table 9. *Ratio of Imported to Domestically Produced Machinery, Korea and Japan*

Sector	*Korea (1985)*	*Japan (1980)*
General machinery	3.04	0.06
Electrical machinery	0.27	0.04

Source Author's calculations based on input-output tables contained in Bank of Korea (various issues), Bank of Japan (various issues)

observed values of total factor productivity growth, which includes any benefits from the posited spillovers among sectors.

Some interactions are not captured by the input-output transactions shown in tables 7 and 8. In particular, the interactions between the producers and *final* purchasers of machines are not given because investment is a final demand. In Korea, the value of imports of nonelectrical machinery was three times that of domestic production (table 9). It is difficult to argue that there were no imported substitutes or that special adaptations to local conditions were quantitatively significant. Even if locally produced equipment was less expensive and conferred some cost reductions for its users, it would have affected only one-quarter of annual general machinery investment as late as 1985

For Japan, the evidence is more ambiguous; Japan imports very little industrial equipment relative to domestic production (table 9). It thus is possible that domestic manufacturers may have designed equipment that was not available from imports and that did increase productivity.

Investment Coordination

The establishment of a domestic steel industry may provide low-cost steel for auto producers, but it does not pay to build a steel mill unless there is an auto industry to use the steel, guaranteeing it a large market and enabling it to realize economies of scale. Conversely, the auto factory may not be built if it is unable to obtain low-priced steel. Investment coordination, which was part of government policy in Japan and Korea, could have benefited both sectors (Pack and Westphal 1986). However, the prices that allow viability could have been obtained in the first place if inputs (steel used in autos) had been available at world prices. Conversely, although the steel sector may benefit from scale economies as its level of production increases, this increase in productivity could have been obtained by exporting steel, as indeed occurred in Japan and Korea. Both countries relied on exports to reduce the need to coordinate investment in sectors where economies of scale were important. Even if one abstracts from the fact that investment coordination only generates benefits where one of the goods is not traded and asks whether Japan and Korea could have ob-

tained significant benefits from such coordination, the data in tables 7, 8, and 9 suggest that the gains from coordination would have occurred largely within the heavy industries, as very few neglected sectors made large purchases from them. Thus any productivity gains from coordination are already reflected in the higher productivity growth of the heavy industries.

Some discussions have noted that Korea's export growth in the 1980s and 1990s has consisted increasingly of products that were manufactured in the promoted sectors, and this has been taken as a sign of success. However, just as the sectoral structure of production would have changed as the use of human and physical capital became more intense, so the structure of exports would have changed. Thus it is necessary to consider the evolving export structure relative to that of other countries. Noland (1997) shows that the export structure of the East Asian countries (including Korea) is not much different than would be expected on the basis of factor endowments. Yoo (1990) compares Korea's exports in the 1980s with those of Taiwan, which did not intensively promote the machinery and chemical industries. In almost all product categories that fall within the HCI complex, Taiwan's export growth was similar to or greater than that of Korea's, suggesting that the growth in exports from these sectors resulted from changing comparative advantage rather than from sectorally targeted intervention.

There are, of course, many individual examples of spectacular success that have led some observers to infer a much more important role for industrial policy, but the aggregate impact of such efforts was limited. Some analysts point to the technological complexity of Korean and Japanese industry, contrasting it favorably with that of Hong Kong. But such complexity is not an end in itself—it must yield higher growth per capita than would have been achieved otherwise. The results reported here fail to measure such an effect. Moreover, the complexity usually considered is that of the manufacturing sector. As the events of late 1997 and 1998 underline, a broader view of the economy, including the capabilities of the financial sector, suggests that Hong Kong was hardly inferior to either of the two industrial giants.

Conclusions

In contrast to other cases of interventionist development strategies, Korea and Japan had vastly better experiences, having generated higher growth rates than many economies that pursued import-substituting industrialization. What accounts for the difference? The simplest explanation is that policies in both countries induced significant competition, whether by holding "contests," as in Japan (Stiglitz 1996), or by linking preferential interest rates and tariffs on imported goods to success in export markets, as in Korea. Firms thus had strong incentives to improve productivity. Other countries that attempted to encourage specific sectors relied on protecting the do-

mestic market and never credibly sought to reduce such protection. Profits and wages were never threatened, and incentives to learn were weak. Countries attempting to extract the benefits from industrial policy that Japan and Korea obtained have to possess not only an exceptionally capable bureaucracy but also the political ability to withdraw benefits from nonperforming firms. Experience in dozens of other countries suggests that these conditions rarely obtain.

Using a variety of assumptions, perhaps as much as 1 percentage point of growth in the manufacturing sector might have been attributable to industrial policy, implying one-third of 1 percent of GDP growth. The implied increase in the national marginal product of capital and labor also may have led to some increase in investment rates in physical and human capital; the magnitude would depend on the unknown elasticity of such investment with respect to the rate of return. Allowing for such secondary effects, the increase in aggregate growth rates induced by industrial policy may have been perhaps half of 1 percent a year, hardly trivial, but not the secret of success. To answer the question posed by the title of this article, industrial policy may have been a minor growth hormone. It seems unlikely that, absent a host of other factors that impinged on both Japan and Korea, the delayed adverse side effects of industrial policy on the banking system were the poison that necessitated very strong antidotes in the late 1990s. But in view of the minor benefits and the potentially adverse effects on the financial sector and the neglected industrial sectors countries should be exceptionally cautious before embarking on such policies.

Notes

I have had helpful conversations on many of the topics of this article with John M. Page, Jr., Joseph Stiglitz, Morris Teubal, and Larry E. Westphal. They do not share all of the views expressed in this article. Research support from the World Bank and the University of Pennsylvania Research Foundation is gratefully acknowledged

1 The estimates for Japan and Korea overstate the correct value for TFP growth, since data on the impact of education on the labor force are omitted. Thus the calculations of TFP growth overestimate the contribution of industrial policy by some unknown amount

2 The estimates are based on Tornqvist indexes. I have used the data given by Yoo (1990) for Korean value added, capital stock, and labor force. Yoo's data on capital stock are derived from a study of Pyo (1988). The data are based on consistent input-output definitions of sectors and include a number of adjustments of the labor force to allow for unpaid workers.

3. Their results, however, are not directly comparable because Dollar and Sokoloff use fixed-weight estimates of factor shares rather than a Tornqvist index and their data on capital, labor, and value added rely on different sources (see footnotes to table 1 in Dollar and Sokoloff 1990)

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Amsden, Alice H. 1989. *Asia's Next Giant: South Korea and Late Industrialization.* New York: Oxford University Press.

Balassa, Bela, and associates. 1982. *Development Strategies in Semi-Industrial Economies.* Baltimore, Md.: Johns Hopkins University Press

Bank of Japan. Various issues *Japanese Economic Statistics* Tokyo

Bank of Korea. Various issues *Monthly Statistical Bulletin* Seoul

Bhagwati, Jagdish. 1978. *Foreign Trade Regimes and Economic Development Anatomy and Consequences of Exchange Control Regimes* Cambridge, Mass · Ballinger

Chenery, Hollis B 1959 "The Interdependence of Investment Decisions " In Moses Abramovitz, ed., *The Allocation of Economic Resources.* Stanford, Calif. Stanford University Press.

Claessens, Stijn, Simeon Djankov, and Larry Lang 1998. "East Asian Corporates: Growth, Financing, and Risks over the Last Decade " Policy Research Working Paper no 2017. Finance, Private Sector, and Infrastructure Network, Economic Policy Unit, World Bank, Washington, D C

Dollar, David, and Kenneth Sokoloff. 1990. "Patterns of Productivity Growth in South Korean Manufacturing Industries, 1963–79." *Journal of Development Economics* 33(October) 309–27

Erzan, Refik, Kiroaki Kuwahara, Serafine Marchese, and Rene Vossenar 1989 "The Profile of Protection in Developing Countries." *UNCTAD Review* 1 29–73

Fry, Maxwell J 1995 "Financial Development in Asia Some Analytical Issues " *Asian-Pacific Economic Literature* 9(May).40–57

Harrigan, James 1995 "Factor Endowments and the International Location of Production Econometric Evidence for the OECD, 1970–85 " *Journal of International Economics* 39(August).123–41.

Itoh, Motoshige, and Kazuharu Kiyono 1988 "Foreign Trade and Direct Investment " In Ryutaro Komiya, Masahiro Okuno, and Kotaro Suzumura, eds., *Industrial Policy of Japan.* Tokyo Academic Press

Johnson, Chalmers. 1982. *MITI and the Japanese Miracle The Growth of Industrial Policy, 1925–75* Stanford, Calif Stanford University Press

Jones, Leroy, and Il Sakong 1980 *Government, Business, and Entrepreneurship in Economic Development The Korean Case.* Cambridge, Mass.. Harvard University Press.

Jorgenson, Dale W , Masahiro Kuroda, and Mieko Nishimizu 1987 "Japan-U S Industry-Level Productivity Comparisons, 1960–79." *Journal of the Japanese and International Economies* 1(March):1–30

Krueger, Anne. 1978. *Foreign Trade Regimes and Economic Development Liberalization Attempts and Consequences* Cambridge, Mass . Balinger

Krueger, Anne, and Baran Tuncer. 1982 "An Empirical Test of the Infant Industry Argument " *American Economic Review* 72(December) 1142–52.

Lall, Sanjaya. 1997. *Learning from the Asian Tigers Studies in Technology and Industrial Policy.* New York St. Martin's Press

Lawrence, Robert 1993 "Japan's Different Trade Regime· An Analysis with Particular Reference to Keiretsu " *Journal of Economic Perspectives* 7(March) 3–20.

Little, Ian M. D. 1982 *Economic Development.* New York. Basic Books.

Little, Ian M. D., Tibor Scitovsky, and Maurice Scott. 1970. *Industry and Trade in Some Developing Countries* New York: Oxford University Press

Mutoh, Hiromichi. 1988. "The Automotive Industry." In Ryutaro Komiya, Masahiro Okuno, and Kotaro Suzumura, eds., *Industrial Policy of Japan.* Tokyo: Academic Press.

Nam, Chong-Hyun. 1990. "Export Promotion Strategy and Economic Development in Korea " In Chris Milner, ed., *Export Promotion Strategies* New York New York University Press.

Nishimizu, Mieko, and John M. Page 1987. "Economic Policies and Productivity Change in Industry: An International Comparison." Development Research Department, World Bank, Washington, D.C. Processed

Noland, Marcus. 1997. "Has Asian Export Performance Been Unique?" *Journal of International Economics* 43(August) 79–101

Pack, Howard. 1988. "Industrialization and Trade." In Hollis B. Chenery and T. N. Srinivasan, eds *Handbook of Development Economics* Amsterdam North-Holland

Pack, Howard, and Larry Westphal. 1986. "Industrial Strategy and Technological Change." *Journal of Development Economics* 22(June)·87–128.

Pyo, Hak-Kil 1988. "Estimates of Capital Stock and Capital-Output Coefficients by Industries for the Republic of Korea." Working Paper 8812 Korea Development Institute, Seoul Processed

Rhee, Yung Whee, Bruce Ross-Larson, and Gary Pursell. 1984 *Korea's Competitive Edge Managing the Entry into World Markets* Baltimore, Md.. Johns Hopkins University Press

Rodrik, Dani 1995 "Getting Interventions Right How South Korea and Taiwan Grew Rich." *Economic Policy* 20(April) 55–107

Rosenberg, Nathan 1976 *Perspectives on Technology* London Cambridge University Press

Sakong, Il 1993 *Korea in the World Economy* Washington, D C · Institute for International Economics

Saxonhouse, Gary 1983 "What Is All This about Industrial Targeting in Japan?" *World Economy* 6(September) 253–73

——— 1993 "What Does Japanese Trade Structure Tell Us about Japanese Trade Policy?" *Journal of Economic Perspectives* 7(Summer)·21–43

Scitovsky, Tibor 1954 "Two Concepts of External Economies." *Journal of Political Economy* 62(April).43–51

Shouda, Yasutoyo 1982 "Effective Rates of Protection in Japan " *Nihon Keizai Kenkyu* 11(March)

Stiglitz, Joseph 1994 "The Role of the State in Financial Markets " *Proceedings of the World Bank Annual Conference on Development Economics 1993.* Washington D.C.· World Bank

——— 1996 "Some Lessons from the East Asian Miracle " *The World Bank Research Observer* 11(2) 151–78

Syrquin, Moshe, and Hollis B Chenery 1989. *Patterns of Development, 1950 to 1983.* World Bank Discussion Paper 41 Washington, D C.

Tanaka, Naoki 1988. "Aluminum Refining Industry " In Ryutaro Komiya, Masahiro Okuno, and Kotaro Suzumura, eds , *Industrial Policy of Japan* Tokyo· Academic Press

Trezise, P. H. 1983. "Industrial Policy Is Not the Major Reason for Japan's Success." *Brookings Review* 3(2).13–18

United Nations. Various issues. *Yearbook of Industrial Statistics* New York.

Westphal, Larry E 1978. "The Republic of Korea's Experience with Export-Led Development" *World Development* 6(March)·347–82

ılf, Charles. 1988. *Markets or Governments· Choosıng between Imperfect Alternatıves* Cambridge, Mass.: MIT Press.

ırld Bank. Varıous years. *World Development Report.* New York Oxford Unıversıty Press.

namura, Kozo. 1986. "Caveat Emptor: The Industrıal Polıcy of Japan." In Paul Krugman, ed., *Strategıc Trade Polıcy and the New Internatıonal Economıcs* Cambrıdge, Mass.· MIT Press.

ıamura, Kozo, and Yasukıchı Yasuba 1985 *The Polıtıcal Economy of Japan.* Stanford, Calıf: Stanford Unıversıty Press.

ıawaki, Hideki. 1988 "The Steel Industry" In Ryutaro Komıya, Masahıro Okuno, and Kotaro Suzumura, eds., *Industrıal Polıcy of Japan* Tokyo Academıc Press

ıazawa, Ippıe. 1988 "The Textıle Industry" In Ryutaro Komıya, Masahıro Okuno, and Kotaro Suzumura, eds, *Industrıal Polıcy of Japan.* Tokyo Academıc Press

ezawa, Yoshie. 1988 "The Shıpbuıldıng Industry" In Ryutaro Komıya, Masahıro Okuno, and Kotaro Suzumura, eds., *Industrıal Polıcy of Japan* Tokyo Academıc Press.

, Jung-ho. 1990. "The Industrıal Polıcy of the 1970s and the Evolutıon of the Manufacturıng Sector ın Korea." Workıng Paper 9017 Korea Development Instıtute, Seoul Processed

Financial Safety Nets: Lessons from Chile

Philip L. Brock

Should governments ever override bank regulators who are attempting to close down insolvent financial institutions? An analysis of Chile's history shows that time after time from the 1850s to the 1980s, prudential banking regulations were abandoned during economic crises when attempts to impose tight solvency standards proved impossible to enforce. Chile's current stringent banking regulations may prove more durable, but mounting financial distress is equally likely to lead the government to adopt policies that prevent bank failure but undermine the authority of regulators.

Bank regulators, including the central bank, are responsible for creating a financial safety net to protect depositors against loss and for enforcing the rules of prudent behavior that are required for a stable financial system. Because safety nets often additionally cover losses to bank owners and borrowers, the support they offer encourages risk-taking by the private sector—an action that may promote financial deepening, but at a high budgetary cost to the government. Poorly designed safety nets may have to be suspended during crises to prevent losses from mounting and to limit the government's liability.

The central aim of government-created institutions and regulations establishing financial safety nets has been to strengthen the financial system's ability to withstand bank runs and severe economic shocks while at the same time shifting risk to the government and promoting financial deepening (an increase in the stock of financial assets). Although it is frequently said that the best safety net is one that results in market participants behaving as if it did not exist, well-designed safety nets *should* alter bank behavior.

This article explores the issues surrounding financial safety nets in small, open economies, drawing on Chile's experience in the 1980s. It argues that a good institutional safety net will balance the benefits of financial deepening with the costs of shifting risk from the private sector to the government. Financial deepening, which is typically measured as an increase in the ratio of bank liabilities to gross domestic product (GDP) or in broad measures of bank assets to GDP, is generally regarded as a

The World Bank Research Observer, vol. 15, no. 1 (February 2000), pp. 69–84

good thing for an economy because it relaxes self-financing constraints and is associated with the increasing monetization of transactions. The components of a good safety net—including lender-of-last-resort facilities, deposit insurance, capital requirements, supervision, and closure and recapitalization rules—must be designed to control the amount of risk borne by the government. Without prudential supervision, the anticipated benefits of financial deepening may be illusory.

The discussion shows how government regulation can lower the cost of lending. The creation of better bankruptcy laws, the granting of special legal powers to bank boards, and the requirement that liquid bank liabilities be backed up by liquid assets are all ways in which prudential regulations can expand the scope for intermediation by banks and other financial institutions. Although prudential regulation does not always accompany the creation of an explicit safety net, it frequently results in government intervention during financial crises, even when the government has no formal role. The article then describes the creation of central banks as an example of an innovation to promote financial deepening and expand intermediation by formally monitoring bank operations and by acting as lender of last resort to ensure the liquidity of the banking system.

The extension of the financial safety net to depositors is a relatively new and untested development in many countries. If the state has a comparative advantage in monitoring banks, deposit insurance has the potential to improve welfare and expand output, but it also may misallocate an economy's resources. For most small, open economies the long-term challenge to prudential regulation of the financial system is to protect depositors in the wake of large macroeconomic shocks.

Characteristics of Prudential Regulation

Banks typically have been subject to a moderate amount of regulation committing them to observe more stringent prudential practices than they otherwise would guarantee depositors. Under this sort of regulation, bank closures generally are triggered by bank runs, and the government remains outside the compensation process as far as depositors and other bank creditors are concerned. Even so, pressures arise to provide aid to debtors during severe economic downturns. An example from Chile illustrates how borrowers who owed money to a financial institution with no explicit safety net were able to generate enough pressure on the government to force it to create an ex post safety net.

Borrower Distress in the 19th Century

During the gold rush of the early 1850s, demand for grain exports from Chile to California and Australia surged. New loans for large landowners who were anxious

to expand production were unavailable, however, because existing mortgage laws were poorly defined in legal terms and because potential lenders lacked reliable information on the holdings, quality, and legal status of the properties involved.

In 1856 the Chilean Congress responded by creating a special mortgage giving clear rights in case of default to the lender. Property registries were set up to make information on sales, mortgages, and censuses readily available, and a state-sponsored mortgage bank, the Caja de Crédito Hipotecario, was established.[1] The property registries helped to mitigate the information problem facing lenders, and the Caja became a monitor that could reduce the costs of lending by adhering to legally mandated collateral requirements, holding a diversified portfolio of loans, and applying economies of scale in monitoring. The new mortgage law gave clear authority to the Caja to enforce bankruptcy proceedings if a landholder fell sufficiently behind in making mortgage payments.

The reform of mortgage laws and the creation of the Caja shifted some risk from landowners to the Caja and to purchasers of its securities. By the end of 1860 more than 5 million pesos of the Caja's securities were in circulation (equal to 10 percent of exports). The financial deepening associated with these agricultural loans propelled the expansion of irrigation and other improvements to the land.

Although the rapid five-year credit expansion permitted landowners to undertake capital improvements on their land, it also exposed them to macroeconomic shocks. When the California and Australian export markets collapsed in the late 1850s, many Chilean landowners were unable to make their mortgage payments. In 1858–59, responding to intense pressure to prevent the foreclosure of landholdings, the government clandestinely funneled about 2 million pesos to landholders (which it skimmed from a 7-million-peso railroad loan that had been financed in the London bond market). This action, which effectively shifted the risk to the taxpayers, was not envisioned by the Caja's institutional structure and was opposed by the foreign investors who had supplied the capital for the railroad loan. In 1860 a new government was installed, and it tried to recover the clandestine loans. The resulting economic contraction of 1861 and 1862, which produced the liquidation of a large number of landholdings, was Chile's first financial crisis and one of the most severe economic contractions of the 19th century (Brock 1992; Fetter 1931).

This example highlights features common to the dynamics of many implicit financial safety nets. The Caja and the accompanying legal reforms were created to promote financial deepening. In theory, prudential lending practices were legally mandated so that the risk to the Caja was carefully controlled. In practice, though, the Caja expanded mortgage lending so quickly that it became overexposed to the risk of a mass default by landowners in the event of a crisis. The threat of massive bankruptcies initially created pressure for a remedial safety net underwritten by the government and then contributed to a severe economic downturn when the properties were finally liquidated.

Governments have devised ways to deal with these kinds of problem by establishing central banks and other financial institutions to supervise banking activity. The creation of the Federal Reserve System in the United States in 1914 is often described as a legislative response to instability in the financial sector.

Similarly, the introduction of central banks into Latin America in the 1920s and 1930s had far-reaching consequences for financial deepening and for incentives to shift risk to the government. The Central Bank of Chile, like that of several other countries in Latin America, was established on the recommendation of a commission headed by Edwin Kemmerer, a professor at Princeton University, who proposed a central bank and a superintendency of banks that would jointly watch over the financial system.

The institutions and laws recommended by the Kemmerer Commission modernized banking rules and made all banks subject to standardized conventions of management and accounting, and they established minimum ratios of capital to assets and of reserves to liabilities. Bank superintendents were created to examine the banks and to enforce banking legislation (the detailed rules and norms issued by the central banks). The superintendencies were located outside the central banks under the aegis of the finance ministries, but they were set up as semipublic institutions and were guaranteed independence from the finance ministries and the rest of the government. The central bank was required to hold gold and foreign exchange reserves sufficient to cover at least 50 percent of its liabilities.

The key policy tool was the rediscount window, which allowed central banks to extend credit only against the collateral of short-term commercial paper (Tamagna 1965). Central banks established one set of discount rates for banks and a slightly higher set of discount rates for the public. The former were intended to foster the role of the central bank as the lender of last resort. The slightly higher, but moderate, public rates were designed to put pressure on banks to lower interest rates on loans and make rates more uniform.

The new legislation and institutions did not create an explicit safety net for banks. They did set up capital requirements, provisions for monitoring financial institutions, and authority to close banks. These measures reduced the costs of financial intermediation and led to financial deepening by decreasing the costs of borrowing.

The legislation in Chile created a guarantee that domestic currency would be convertible into foreign exchange at a fixed exchange rate. This, in combination with the newly instituted oversight of the banking system, created strong incentives for foreign investment. In essence, the fixed exchange rate with fiscal oversight of the banks created an implicit guarantee of the banks' foreign exchange liabilities, thereby shifting the risk to the government. As more investment was intermediated through the financial system, financial depth increased.

The Emergence of the Safety Net

The conditions that led to the creation of the safety net in Chile were similar to those that have become familiar in more recent financial crises in the region. In the late 1920s Chile's banks were financially stable. They operated with capital-to-asset ratios of 20 to 30 percent; they were rigorously monitored by the Superintendency of Banks; and the mechanisms for closing troubled banks were enforced. By normal standards, there was no need for a safety net for the banks. Indeed, the whole apparatus lowered the cost of financial intermediation, so that depositors required a smaller premium, borrowers paid lower loan rates, and foreigners were assured that their loans would be repaid in foreign currency. All of these effects enhanced financial deepening between 1925 and 1929.

Following the creation of the Central Bank and the passage of accompanying banking legislation in 1925, the economy grew at an average rate of about 10 percent a year from 1926 to 1929. With the start of the Great Depression, Chile's GDP fell 11 percent in 1930, 17 percent in 1931, and 27 percent in 1932—a cumulative decline of more than 50 percent in three years. As the economy spiraled downward, the Superintendency of Banks and the Central Bank raised deposit rates in several incremental steps up to the third quarter of 1931. The superintendency also took strong steps to force banks to adjust to the deterioration in the quality of their loans and investments, including stepping up inspections of loan portfolios, requiring bonds and real estate investments to be priced to reflect the downturn in prices, enforcing provisioning against possible loan losses, and setting penalty interest rates on nonperforming loans. The immediate result was a rise in nonperforming loans, which climbed from 4.3 percent of total bank assets in 1929 to 15.3 percent by the end of 1931. As a result of the economic and financial deterioration, 4 of the 22 banks in the country were forced to close in 1930 and 1931.

Through mid-1931 the economy's adjustment followed the rules of the gold standard—outflows of gold produced a contraction in the money supply. Between January 1930 and July 1931, the Central Bank lost more than half its gold holdings but still had a reserve of 72 percent against its demandable liabilities (notes and deposits). From July 1931 to the beginning of June 1932, economic policy began to deviate from the automatic adjustment of the money supply to changes in the Central Bank's gold holdings. Exchange controls were announced on July 30, 1931, to protect the remaining gold reserves, following an earlier default on Chile's foreign debt.

In mid-1932, amid growing civil opposition to the government's largely inactive economic policies, the military staged a coup and established a socialist republic, which lasted only 100 days. During this time, the president, who was given vast powers to intervene in the economy, introduced trade quotas, import licenses, and a jobs program for the unemployed. The government also devalued the currency 70 percent and declared a three-day bank holiday (June 6–8, 1932), during which bank

deposits were frozen and foreign funds in the banking system were declared the property of the state. In an effort to bail out private debtors, a moratorium on the repayment of debts was announced (Ellsworth 1945).

In the next six months, successive governments increased domestic credit by about 800 million pesos, equal to about 70 percent of the banking system's total loan portfolio (Chamorro 1985). Recovery was rapid; in the first year real GDP rose 16 percent, due in part to an improved external environment. The government also announced a 10-year tax holiday for all new commercial and residential construction under way by the end of 1935, creating a building boom that succeeded in raising depressed property prices. The 30 percent price increase between June 1932 and June 1933 provided debt relief to enterprises that had borrowed funds, while interest rates were controlled at low rates and penalty interest rates were rescinded. In 1933–34 the real value of debt was reduced substantially, and nonperforming loans declined rapidly.

The main lesson here is that orthodox governments frequently respond to a financial crisis by initially affirming that there is no safety net, only to announce later a series of emergency rescue measures that create one in response to pressures generated by widespread financial distress among borrowers. Did the new Central Bank and bank regulations lower the cost of financial intermediation too much? "Too much" refers in a narrow sense to the government's ability to make good on its fixed exchange rate guarantee. In a broader sense it refers to the government's desire to insure against large macroeconomic shocks. In Chile, as in much of Latin America during the 1930s, the answer was that governments could not guarantee the banking system under the rules of the gold standard. The ad hoc bank safety nets created at the start of the Great Depression saved the banks at the cost of capital levies on depositors, moratoriums on foreign debt service, and emergency loans from the central banks.

Financial Liberalization

The attempts at financial liberalization that began in the Southern Cone of South America in the 1970s and spread throughout the world during the succeeding two decades have not been painless. Liberalizing countries eliminated targeted credit programs and lifted interest rate controls. In addition, new banks were allowed to begin operations. However, in most countries the existing financial regulatory structures were not designed to require that banks be evaluated when they reshaped their portfolios from government-directed loans toward much higher yield loans to companies and consumers. Monitoring their borrowers has proved to be difficult for banks, and deficient internal bank controls on both initial and ongoing loan evaluations have allowed banks to lend even in circumstances in which repayment seems unlikely

Government bank examiners have been equally incompetent; many of the best examiners have been hired away by banks at high salaries, and those who remained have been too few and too powerless to meet the demands placed on them. During liberalization, bank capital often has been inadequate, and even published capital-to-asset ratios frequently have been overstated by officials anxious to conceal bad loans. Bank managers may even resort to accounting gimmicks such as double-gearing, whereby a bank lends, say, $100 to a firm within its group so that the firm can buy $100 of the bank's stock. The bank's reported capital rises by $100, but the group's capital remains the same.

Because financial liberalization frequently results in increased risk, governments typically guarantee funds lent to banks by domestic and foreign lenders who otherwise would hesitate to invest in banks that might fail. The Chilean government's rescue of Banco Osorno in early 1977, for example, saved foreign creditors from losses that would have put an end to the capital inflows that were helping to fuel the economic recovery. The implicit government guarantees associated with that rescue meant that bank spreads (that is, the difference between loan rates and deposit rates) were kept much lower than they would have been otherwise, attracting more borrowers and encouraging financial deepening. The ratio of private sector domestic credit to GDP rose from 8.8 percent in 1977 to 39.3 percent in 1981 (Brock 1996). But in the context of poor monitoring capacity and low bank capital, the implicit guarantees meant that the government was assuming excessive risk.

The apparent initial success of Chile's liberalization was held together on the strength of an implicit government guarantee to protect depositors and other bank creditors. The true extent of the safety net was unveiled only as the financial crisis began. Because governments virtually always step in to save banks from failing, even when there is no explicit provision for a safety net, financial structures should be designed to take into account past experience with these ex post safety nets. The example of Chile illustrates the point and demonstrates that once the financial safety net has been thrown out to save banks, untangling it may be complex and time-consuming.

The Implicit Safety Net in Chile

In Chile, as in many other countries, the financial crisis initiated an exchange rate crisis. Throughout 1981 and the first six months of 1982, the financial sector continued to deteriorate. In June 1982 the government devalued the currency and instituted a preferential exchange rate for dollar debtors—a subsidy that would amount cumulatively to about $3.4 billion by the end of 1987—to reinforce the government's implicit guarantee that the fixed exchange rate would be a permanent anchor for the economy. In July the central bank began to buy part of the banks' bad loan portfolios at face value and agreed to buy back the rest of the loans over a three-to-five-year

period. This action improved the banks' balance sheets by replacing nonperforming assets with a Central Bank bond, but because the bond paid no interest and was not transferable, no resources were transferred to the banks.

Regulatory forbearance also played a role in the safety net. In September 1982 the Superintendency of Banks allowed banks to use the June 30 exchange rate in calculating the peso value of their dollar liabilities. This represented a 35 percent underestimation of the value of dollar liabilities by the end of 1982, but it prevented several banks from violating minimum regulatory capital ratios. Banks originally were told to provision against these losses by the end of 1982, but the deadline later was extended to the end of 1986. In addition, in October 1982 the authorities extended the time limit for declaring a loan nonperforming from 30 days following failure to make a loan payment to 90 days.

The intervention in the flagship banks of the two largest economic conglomerates on January 13, 1983, also stabilized the financial system by terminating the creation of shell companies (that is, companies that existed only on paper) and by blocking other measures such as the use of offshore subsidiaries to evade prudential regulations. This step brought more than 50 percent of the financial system's assets and liabilities under the direct control of the government and provided explicit government backing of the liabilities of the remaining institutions. It also gave the government control of several of the largest firms in the economy.

In a major departure from previous policy, the government announced that all existing bank debt, both external and domestic, would now be guaranteed by the government. Although the new law created a huge contingent liability for the government associated with bank insolvencies, it prevented a run on the banks and, by converting external debt into sovereign debt, averted the forced liquidation of banks by foreign creditors.

By early 1983 it had become clear that many loans would not be repaid at the contracted terms. In April of that year, in response to the growing threat of a debtor revolt, the government announced a restructuring plan to reschedule 30 percent of an eligible company's debts for 11 years, with a 1-year grace period for interest and a 5-year grace period for principal. The operation created a cash flow subsidy to the banks that amounted to 7 percent of the amount of the restructured loans. The last major stabilization measure, announced in June 1983, protected holders of home mortgages by rescheduling installments unpaid since 1981.

Modifying the Guarantees

One year after the devaluation, all the safety net measures were in place. Three of the programs were modified substantially the next year. In February 1984 the government expanded its repurchase program and shortly thereafter extended the produc-

tive debtor plan and home mortgage refinancings. None of these measures entailed restructuring the financial system; the largest banks and enterprises were still in limbo, under temporary government control but without any plan to restore private ownership.

In 1985 a new finance minister raised the argument with foreign creditors that the Chilean government had provided $3.4 billion to date to help borrowers repay their dollar loans and that it was time for the foreign creditor banks to bear their share of the losses. For their part, the creditor banks were pressing Chile to renew its guarantee on bank debt, much of which was due to mature during 1985–87. In subsequent negotiations the government agreed to the renewal, and the banks' creditors gave tacit approval to debt buyback and debt-equity conversion mechanisms. Between mid-1985 and mid-1987 about $3.2 billion in bank debt was written down or converted into equity by these mechanisms at market discounts of about 30 percent. These transactions were essential for recapitalizing Chile's banks.

Between 1982 and 1984, 17 private banks had exchanged bad loans for Central Bank bonds. In January 1985 a "popular capitalism" law authorized the Central Bank to capitalize a portion of its loans to the five banks over which the government had assumed control. The law resulted in the recapitalization of four of the five banks (the fifth merged with another bank). Existing stockholders had first claim on stock purchases. The remainder was given to the Chilean Development Corporation (CORFO) and then sold in small amounts to individual investors.

In 1989 concern regarding the possible fiscal manipulation by future governments of banks' ability to repay led the outgoing military government to redraft the agreements between the banks and the Central Bank. In place of a fixed obligation, the new law created a debt obligation that was equal in nominal value to the fixed obligation but that had no fixed timetable for payment. The new contracts were protected by law against unilateral changes by the Central Bank. In the case of the five heavily indebted banks, there was effectively no date by which the subordinated debt would ever be repaid.

When the newly elected democratic government took power in 1990, the total value of the subordinated debt of 11 banks was about $3.3 billion, while the value of paid-in capital was only $1.3 billion. The new government wanted changes that would eliminate the indeterminacy of the resolution of the debt problem. Despite renewed efforts in 1992–93, the government made no appreciable progress in drafting a replacement for the subordinated debt law that would be acceptable to the banks. In 1995, after a series of disputes, an agreement was drafted stipulating that the banks would repay part of the debt in exchange for forgiveness of the remainder. During the negotiations the president of the Central Bank, who argued that the banks should be forced to assume a greater portion of their losses, resigned. Ultimately, the Central Bank wrote off approximately $2 billion to recapitalize the banks.

Deposit Insurance

All financial systems have an implicit as well as an explicit safety net. Until the late 1980s, depositors generally were not included in the explicit safety net. Indeed, they generally became part of the ex post measures to prevent bank failures. In Argentina and Brazil in the 1980s deposits were frozen and then reduced in real value by high inflation as a way of improving bank solvency. Even in Chile, where deposits were guaranteed in January 1983, interest rates on deposits were kept artificially low by a combination of capital controls and Central Bank suggestions regarding the appropriate rates to pay. Since 1981 the number of countries with explicit deposit insurance programs has tripled, from 15 to 45. These programs represent a significant, but untested, institutional innovation in the adopting countries. Providing a safety net for depositors lowers the cost of deposits to banks by moving contracting problems associated with asymmetric information to the government insuring agency. As with other contractual innovations, deposit insurance leads to financial deepening and a shift in financial risk. In this way it is no different from loan contracts that banks negotiate with their borrowers. For loan contracts to remain viable in the long run, banks must be able to restrict the ability of borrowers to shift risk to the banks. Similarly, if deposit insurance is to remain viable, the government must restrict banks' ability to shift risk to the insurance agency and to the taxpayers. The best example of the perverse dynamics of risk-shifting coupled with financial deepening in the context of deposit insurance comes from the crisis in the U.S. savings and loan industry in the 1980s.

As a result of high interest rates paid to attract depositors during the late 1970s and early 1980s, the net worth of many savings and loan institutions in the United States deteriorated. The Federal Reserve Board originally responded to the problem by imposing interest rate ceilings on deposits, in effect taxing primarily small depositors to keep the savings and loans solvent. When the creation of new financial instruments, such as mutual funds, caused a flight of funds away from the savings and loans, Congress passed legislation to help the industry by deregulating its asset powers, permitting savings and loans to issue credit cards and to make consumer loans up to 30 percent of assets; to make commercial real estate loans up to 40 percent of assets; to offer commercial loans up to 11 percent of assets, and to take direct equity positions up to 3 percent of assets. State chartering authorities, especially in California, Florida, and Texas, provided even wider asset powers to thrifts in what appears to have been regulatory competition to keep thrifts from adopting national charters. During 1983 and 1984 the resulting asset growth of savings and loans (19 percent a year) far outstripped the asset growth of commercial banks (7 percent a year). Romer and Weingast (1990) argue that the fast-growing thrifts were gambling for resurrection and were able to do so because Congress forced the Federal Savings and Loan Insurance Corporation (FSLIC) to pursue a policy of forbearance that prevented the closure of undercapitalized savings and loans.

Although deregulation expanded bank lending, it also shifted excessive risk to the government. The government, for its part, failed to enforce capital standards and curtailed inspections of savings and loans during 1984–85. Although the formal safety net (the funds of the FSLIC) was modest, the informal safety net (the funds of the U.S. Treasury) was almost unlimited. Depositors were not worried about risk, and indeed, fast-growing thrifts only needed to offer small premiums on certificates of deposit to generate large deposit inflows.

The ex post safety net to protect depositors against loss involved shutting down more than 700 savings and loans with assets of $400 billion. The Resolution Trust Corporation (RTC), which directly or indirectly employed up to 20,000 people during a three-year period, either liquidated the assets it acquired from failed thrifts or reorganized existing institutions. The cleanup operations of the RTC involved losses to taxpayers of approximately $200 billion.

Prudential Regulation and Aggregate Economic Shocks

Much of the discussion on safety nets is based on the premise that bank-specific risk poses the major challenge for the design of bank capital, supervision, and closure mechanisms. Much of the risk faced by banks in Latin America and elsewhere, however, is aggregate rather than bank specific. Aggregate risk is observable by all agents in an economy: everyone knows when and by how much the price of copper or coffee changes or by how much the London interbank offer rate moves. Aggregate risk need not be borne by banks and other financial intermediaries. Indeed, being forced to bear it may interfere with intermediaries' incentives. Enforcing penalty loan rates on firms, changing ownership of the banks, or ordering the liquidation of assets may be a socially inefficient response to a negative aggregate shock. Such actions in the aftermath of an aggregate shock penalize competent owners and managers and may result in inferior utilization of an economy's assets.

The incentive problem posed by aggregate risk to the financial system is similar to that associated with government disaster relief programs. By providing insurance against catastrophes such as floods and hurricanes, the government offers homeowners an incentive to buy property in high-risk areas. To prevent this outcome, the government should implement tough rules that prevent homebuilders from building too close to the ocean or using construction materials that are too flimsy. Governments should similarly provide a safety net for the financial system in the event of a catastrophic shock to the economy. But the size of the bailout associated with the safety net will be affected by the government's ability to control excessive risk-taking in the years before a shock. It is in the good years before a shock occurs (like the calm years with no hurricanes) that government agencies should be given the incentives to strictly enforce conservative capital, monitoring, and closure policies.

Capital

Risk-adjusted capital requirements can play an important role in the design of a financial safety net *provided* that bank monitoring and closure policies are adequate Without good monitoring, banks have an incentive to engage in accounting tricks that artificially raise the book value of their capital, such as the double-gearing discussed earlier. This technique increases the bank's capital for regulatory purposes without increasing the economic group's real capital.

Banks also have an incentive to underreport and underprovision bad loans. Adequate provisioning is crucial for the long-term viability of risk-adjusted capital requirements. Underprovisioning keeps the return to bank equity artificially high, and the cumulative excess flow of dividends destroys market value capital in a process that amounts to looting the bank. Looting of this sort is easy when an increase in the value of implicit or explicit deposit guarantees automatically offsets the decline in the value of the bank's capital.

Monitoring

Evaluating the true net worth of banks presents technical problems. For example, regulators must decide when loans that are current should be classified as doubtful. When the collateral value of the real estate securing a $30 million mortgage falls from $35 million to $25 million, the loan may stay current for a while, but eventual default is almost inevitable. Regulatory enforcement is especially difficult when the fall in collateral is perceived to result from an external shock rather than from reckless lending practices. Given the possibility of disagreement over market values during a collapse in the real estate market, even market value accounting permits long delays in recognizing losses. A two-year delay in forcing the recognition of losses while the real estate market is adjusting downward, for example, can create the incentives and provide sufficient time for even good bankers to gamble away a bank's future by taking on excessive risk.

Closure Policy

Closure policy is the Achilles' heel of any explicit or implicit government safety net. The inability to close failing banks permits equity holders to roll over loan losses and engage in other risky lending practices, thereby bidding deposits away from other institutions and transmitting incentives for risky lending to the rest of the financial system. Just a few banks operating in this way during good times can weaken the whole system's ability to withstand a large aggregate shock.

Both technical and political reasons explain why closure policy is such a thorny issue for bank regulation. From a technical standpoint, bank liquidation generally is

undertaken only as a last resort to avoid the loss of the bank's value as an ongoing operation. In between forbearance and liquidation is a wide range of possibilities, including voluntary recapitalization by the bank's owners, cash-assisted acquisition by another bank, temporary administration by a government workout agency, and forced capital levies on depositors.

From a political standpoint, allowing a bank to fail not only will incur the wrath of uninsured lenders to the bank but also will run counter to the interests of politicians who depend on bank owners for political support. The problem is to create agreements before a bank failure that make it difficult afterward to renege on the no-bailout position. Because the impending failure of a large bank may disrupt the payments system, the various intervention options must be spelled out in some detail in advance to facilitate the bank resolution process without resort to a political rescue of the bank's owners. About one-fourth of the 1986 Chilean bank law, for example, is devoted to the precise specification of alternative closure and recapitalization mechanisms for banks.

A more subtle political problem involves interagency cooperation in bank interventions and closures. The banking regulatory agency has responsibility for determining whether intervention is called for; the deposit insurance agency must be willing to provide the necessary funds to resolve the insurance problem; and the central bank must provide liquidity while the intervention is resolved. Lack of cooperation by any one of these agencies can derail the closure or recapitalization process. Underfunding the deposit insurance agency, for example, may cause that agency to push for the delay of a bank intervention that would deplete the agency's funds. One possible solution is to create a truly independent central bank that would have the incentive to be tough on banks and the deposit insurance agency in order to guard its capital. This also implies, of course, that the central bank may be unwilling to provide emergency liquidity if it cannot be assured that it will recover those funds eventually.

When to Overrule Prudential Regulators

Various institutional arrangements for prudential regulation do tolerably well in preventing the shifting of risk to the government during normal times. Several writers, including Benston and others (1986) and Calomiris (1997), have advocated the use of subordinated debt requirements to counter the incentive to underprovision bad loans. Subordinated debt creates creditors who can put additional pressure on bank owners and managers to take adequate measures to cover such loans. The big problems arise in the aftermath of a negative aggregate shock when borrowers' debts become impossible to service and banks' capital and reserves become inadequate to cover probable losses. In all these cases it is unlikely that subordinated debt will

prevent bank owners and managers from increasing the bank's risk, since subordinated debt holders may have been wiped out along with the shareholders, with both groups willing to gamble by having the bank take on more risk.

In virtually any small economy, a negative macroeconomic shock will overwhelm a well-designed safety net. Imposing penalty interest rates and shutting down banks, which is the correct course of action given the incentives of bank regulators, will in fact be counterproductive. Chile's decision in the early 1930s to stand by the gold standard and to enforce prudential bank regulation eventually resulted in a military coup.

Regulators should be given strong incentives to control risky behavior that could cause an expensive bailout when a shock occurs. But government, too, has an incentive to avert a crisis by assuming the remaining macroeconomic risk borne by banks. There are several ways to do this: one is for the government to self-insure by accumulating a large fund of liquid resources. Another is to secure international lines of credit; Argentina followed this approach. A third is to rescue the banks using government debt but to require repayment (over many years, if necessary) if a bank wishes to expand into new areas. This was Chile's solution in the 1980s and 1990s.

Authorities must take into account the costs of establishing prudential regulation and safety net institutions that are immune to political pressures. Although these regulations and institutions minimize the private sector's incentives to take risks in good years, they may prove too rigid when a macroeconomic shock hits. If no one in the government can overrule the central bank, and the central bank is intent on enforcing prudential regulations and on guarding its capital, it may take a revolution (such as occurred in Chile in 1932) to put a more extensive safety net in place Conservative prudential regulation and enforcing agencies therefore should be insulated from political pressure, but not so much that true economic catastrophes are made worse by the actions of those agencies.

Because catastrophes such as the Great Depression and the 1980s depression in Latin America are rare events, one may question whether it is worth preparing for them. Perhaps it is enough to put strong prudential regulations and safety nets into place that control moral hazard during normal times. However, the costs associated with not knowing when to relax prudential standards or how to implement rescue packages are quite large (Caprio and Klingebiel 1996)—although few studies have explored the actual transitions from standard prudential regulation to the ex post safety nets generally associated with financial crises.

Conclusion

Safety nets are the result of a government decision to assume risk that otherwise would be borne by depositors and shareholders. Society often benefits from the ex-

ansion of banks and bank lending associated with this protection, but excessive nancial expansion can take place if the government does not limit the amount of isk it assumes (either explicitly or implicitly). The threat of macroeconomic shocks nakes design of prudential regulation difficult in small economies. Because governnents generally assume responsibility for some macroeconomic risk, the private secor may be implicitly encouraged to undertake activities that are especially exposed o such risks. These considerations work in the direction of a tough regulatory framework, but a lack of flexibility may worsen the economy's response to an especially evere macroeconomic shock.

For most small, open economies, a desirable level of financial activity would be npossible without some assurance of government guarantees to depositors and banks. inancial deepening often accompanies strengthened guarantees. These government uarantees shift risk toward the government and imply that at some point the govrnment may need to throw out a safety net that involves government resources. Vise governments in these small economies will create strong supervision to limit ie size of the safety net but will also recognize the value of encouraging financial eepening to promote domestic lending and economic growth.

Notes

Philip L Brock is associate professor at the University of Washington, Seattle. The author thanks Jerry Caprio, Augusto de la Torre, Idanna Kaplan, Daniel Lederman, Yira Mascaro, and Armando Razo for comments on earlier versions of the paper

1 The Caja was authorized to lend up to 50 percent of the assessed value of the real estate and was given legal precedence in the collection of its loans In exchange for the mortgages, the Caja issued letters of credit with maturities of 21–25 years and coupon rates of 5–8 percent Borrowers could then sell the letters on the secondary market for cash Landholders who sold the letters on the secondary market received a higher price than without Caja backing because purchasers demanded a smaller insurance premium, knowing that the Caja, via its capital and reserves, was the primary insurer of the mortgages

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems

Benston, George J., Robert A Eisenbeis, Paul M Horvitz, Edward J Kane, and George G Kaufman. 1986 *Perspectives on Safe and Sound Banking Past, Present, and Future* Cambridge, Mass MIT Press

Brock, Philip L 1992 "The Government-Sponsored Securitization of Mortgages in Nineteenth Century Chile The Macroeconomic Consequences of the Caja de Crédito Hipotecario, 1855–1880" University of Washington, Department of Economics, Seattle Unpublished paper

———. 1996. "High Real Interest Rates and Banking Crises in an Open Economy A Case Study of Chile, 1975–1983." University of Washington, Department of Economics, Seattle Processed

Calomiris, Charles W. 1997. *The Postmodern Bank Safety Net: Lessons from Developed and Developing Economies.* Washington, D.C.: American Enterprise Institute Press.

Caprio, Gerard, Jr., and Daniela Klingebiel. 1996. "Bank Insolvency: Bad Luck, Bad Policy, or Bad Banking?" World Bank, Development Research Group, Washington, D.C. Processed.

Chamorro, Mario. 1985. "El sistema bancario comercial chileno y la crisis de los años treinta" Superintendencia de Bancos e Instituciones Financieras, Departamento de Estudios, Santiago, Chile. November. Processed.

Ellsworth, P. T. 1945. *Chile: An Economy in Transition.* New York: Westport.

Fetter, Frank W 1931. *Monetary Inflation in Chile.* Princeton, N.J.: Princeton University Press.

Romer, Thomas, and Barry R. Weingast. 1990. "Political Foundations of the Thrift Debacle." Working Papers in Economics E-90-22. Hoover Institution, Stanford University, Stanford, Calif Processed

Tamagna, Frank. 1965. *Central Banking in Latin America.* Mexico: Centro de Estudios Monetarios Latinoamericanos.

Lessons in Structuring Derivatives Exchanges

George Tsetsekos • Panos Varangis

The global deregulation of financial markets has created new investment opportunities, which in turn require the development of new instruments to deal with the increased risks. Institutional investors who are actively engaged in industrial and emerging markets need to hedge their risks from these cross-border transactions. Agents in liberalized market economies who are exposed to volatile commodity price and interest rate changes require appropriate hedging products to deal with them. And the economic expansion in emerging economies demands that corporations find better ways to manage financial and commodity risks. The instruments that allow market participants to manage risk are known as derivatives because they represent contracts whose payoff at expiration is determined by the price of the underlying asset—a currency, an interest rate, a commodity, or a stock. Derivatives are traded in organized exchanges or over the counter by derivatives dealers.

Since the mid-1980s the number of derivatives exchanges operating in both industrial and emerging-market economies has increased substantially. What benefits do these exchanges provide to investors and to the home country? Are they a good idea? Emerging markets can capture important benefits, including the ability to transfer risks, enhance public information, and lower transaction costs, but the success of a derivatives exchange depends on the soundness of the foundations on which it is built, the structure that is adopted, and the products that are traded.

Since the mid-1970s interest and exchange rates and prices of primary commodities have fluctuated widely. Major exchange rates—both nominal and real—have varied since the move to the floating-rate system in 1973. Commodity prices have been even more volatile, with large shifts in supply and demand for individual commodities. Unanticipated changes in exchange rates, interest rates, and commodity prices introduce risks that cannot be ignored.

Financial price risk has important implications for both the private and public sectors. Price fluctuations not only affect business profits but can also affect a firm's

The World Bank Research Observer, vol 15, no 1 (February 2000), pp 85–98

survival (Smith, Smithson, and Wilford 1990). Changes in exchange rates intensify competition from other countries. Commodity price fluctuations result in changes in input prices and costs. Similarly, changes in interest rates can lead to financial distress as borrowing costs increase. Governments are also affected by volatile markets, particularly in developing countries, where commodities exports account for a large share of total exports, affecting tax revenues. Moreover, a significant part of the external debt of developing countries carries variable interest rates and often is denominated in currencies that do not match the currency composition of their exports

During the 1980s and 1990s financial markets responded to price volatility by introducing a vast range of instruments for managing price risks.[1] These instruments, which are called derivatives, represent contracts whose payoff at expiration is determined by the price of the underlying asset (a specific currency, interest rate, commodity, stock price, or stock price index). Examples of derivatives include forward, futures, and option contracts, and swaps (see the definitions in the appendix). Derivatives are traded both in organized exchanges and over the counter (OTC) by derivatives dealers. Approximately 75 derivatives exchanges are in operation worldwide, most of them in industrial countries, although a growing number of emerging economies plan to establish their own derivatives exchanges. The enormous levels of global trading for both exchange-traded and over-the-counter products (Abken 1994, Becketti 1993; Remolona 1992; Stout 1996) have attracted the attention of academic literature and the popular press, but relatively little has been published about the internal organization and structure of derivatives exchanges.

A derivatives exchange can be defined as a trading forum or a system that links a central marketplace (such as a trading floor or an electronic trading system), where all those with buying and selling interests in a product designed to permit the shifting of risk can meet, with a mechanism (such as a clearinghouse) for intermediating, validating, and enhancing the credit of anonymous counterparts. Elements of exchange design accordingly need to address issues related to the facilitation of trades and the reduction of credit risk in transactions among market participants.

The primary function of a derivatives exchange is to facilitate the transfer of risk among economic agents by offering mechanisms for liquidity and price discovery. *Liquidity* refers to the ability to buy and sell large volumes of derivatives contracts in a relatively short period of time without affecting prices. If risks are to be transferred efficiently, there must be a large number of agents ready to buy or sell. In other words, there must be agents who desire to reduce risk (that is, to obtain price insurance) and agents who want to accept risk (that is, to provide price insurance). The exchange brings together a large number of participants, making quick transactions at high volumes feasible. *Price discovery* refers to the establishment of a price for an asset. The exchange collects and provides immediate information regarding the price of an asset as traders continuously offer to buy or sell. Participants outside the exchange can easily obtain this information and find what the price for the asset is at any moment.

A critical element in the efficient transfer of risk among market participants is the microstructure of the derivatives exchange, which includes regulatory oversight, trading systems, settlement and clearing procedures, membership of the exchange, ownership of the clearinghouse, and products traded. An examination of the microstructure of derivatives exchanges around the world could yield recommendations for improvements in exchanges, as well as data for analyzing the relationships and linkages between critical exchange parameters and the broad economic environment.

In a well-designed derivatives exchange, resources are efficiently allocated, and risk-sharing arrangements are optimal. Merton and Bodie (1995) developed a framework that is applicable to derivatives exchanges and helps explain the institutional form of the exchange and the particular features of its microstructure. Perhaps the largest single influence has been the evolution of technology, such as communications, computer software, and electronic information systems (Chapman 1994; Stout 1996). The functions of a derivatives exchange may be performed under various organizational arrangements. For example, an American and a European derivatives exchange may have different arrangements for clearing and settling transactions, even though both exchanges may trade similar contracts.

This article provides information on key elements and characteristics of the exchange microstructure and identifies differences in organizational structure among derivatives exchanges. The information and analysis should be of particular interest to emerging economies seeking to develop a derivatives exchange.

Why Establish a Derivatives Exchange?

Economic reasons and national pride motivate the establishment of derivatives exchanges. A country's financial infrastructure is enhanced by the links among hedgers, speculators, and cash markets. A derivatives exchange can improve the allocation of resources, maintain efficient pricing and information flows, and act as a conduit for the transfer of risk within a country—and even across countries. Additionally, because derivatives exchanges make more information publicly available, credit systems and capital markets are more responsive, with uniform repayment regulations and market surveillance; transaction costs are lower; forward prices are more accurate; and resources are better allocated (Chang, Kaplan, and Knapp 1999, Peck 1985).

Derivatives allow risk-averse market participants (such as banks, farmers, processors, and traders) to offset risk among themselves or transfer it to other market participants willing to accept the risk-return ratio. In the process, derivatives attract additional participants who in turn increase the volume of transactions, thus contributing to the creation of a liquid market.

By definition, derivatives contracts traded at an exchange tend to be standardized; a clearinghouse guarantees transactions between parties, acting as buyer to all sellers

and seller to all buyers. By requiring that buyers and sellers deposit funds (margins) as security for their transactions and by adjusting these margins to reflect changes in market prices, the clearinghouse substantially reduces, or even eliminates, the performance (counterpart) risk among transacting parties.

In many emerging markets the price for an asset is difficult to determine. An active derivatives exchange plays an important role in facilitating an efficient determination of prices in the underlying cash (or spot) market by providing improved and transparent information on both current and future prices for an asset. For example, in commodity markets spot prices are often pegged to futures prices because the futures market provides excellent pricing information for the underlying product (for security markets, see Scott 1992). Prices on derivatives markets reflect anticipated supply and demand, and derivatives markets thus enhance the ability of market participants to make decisions about future production, processing, and trade.

The exchange writes the specifications for contracts traded, setting standards for grading, measurement, methods of transfer, times of delivery, and contractual obligations. The standardization makes these contracts conducive to centralized trading on an exchange. The ease with which transactions can be executed and positions opened and closed stimulates high trading volumes and higher usage than does the spot market. This ease of execution opens the exchange to almost anyone who wishes to trade, whether involved in the physical trade (spot market) or not, and contributes to reducing transaction costs.

Derivatives transactions using exchanges abroad or OTC markets can expose individual firms and financial institutions to greater risks, however. The increased exposure stems in part from the same source as the benefits: the ability to transfer risks easily means that some market participants are exposed to greater risk than would otherwise have been the case. In countries with weak accounting, auditing, and disclosure regulations, derivatives can be abused, which can potentially destabilize markets (Garber 1998).

Should emerging economies establish their own derivatives exchanges? The two key benefits from establishing a domestic derivatives exchange are improved price discovery and a higher correlation between the prices of derivatives products and cash prices. In this respect domestic derivatives markets may be more closely correlated with local cash market conditions. These benefits should be compared with the risks and benefits of using already established derivatives exchanges that are more liquid and that stress prudential regulation. Higher transaction volumes usually lower transaction costs, which can outweigh the benefit of having a derivatives contract closely associated with local market conditions. The major risks when using established exchanges are the lower correlation between exchange contract prices and local cash prices, and the risk associated with fluctuating exchange rates.

The establishment of derivatives exchanges in emerging markets requires certain preconditions, including well-functioning cash markets, a large number of traders

and speculators, a legal structure that includes a system of property rights and enforceable contracts, well-functioning credit institutions, the support of the government and policymakers, adequate financial resources (particularly for the clearinghouse), and the absence of competing derivatives products and exchanges (Leuthold 1992).

The Microstructure of Derivatives Exchanges

The success of a derivatives exchange is determined to a large extent by the soundness of the foundation on which it is built. Thus, the design of the formal structures and systems created to ensure orderly trading and execution of exchange transactions is crucial. The microstructure may become a form of competitive advantage to the degree that it motivates, facilitates, or enables price discovery and eliminates asymmetric information.

The microstructure of derivatives exchanges is important for several reasons. It provides insights into the workings of the market through an examination of the features of the exchange and of the linkages that allow a disciplined flow of orders ınd execution of transactions, and it plays an important role in creating and dissemiıating market information. Both trading activity and price determination are sensiıve to institutional arrangements. Moreover, the microstructure ensures the smooth ·xecution of transactions. Finally, market arrangements have implications for the ong-term properties of derivatives contracts and the returns on underlying assets. mportant elements of the microstructure include trading mechanisms, clearing arangements, the regulatory structure, and the choice of derivatives products to be raded.

Trading Mechanisms and Clearing Arrangements

Jepending on the automation and sophistication of the market, trades can be ex·cuted on the floor (open outcry) or through electronic trading (Domowitz 1995: Gennotte and Leland 1994). Increasingly, electronic trading is preferred (Chang, Kaplan, and Knapp 1999; Pirrong 1994).

Clearing trades typically involves processing of transactions, including reporting o transacting parties; confirmation of trades; and matching of orders. The function of the clearinghouse is to eliminate or reduce the counterpart credit (performance) isk by standardizing and simplifying the processing of transactions. The clearingıouse matches all purchases and sales on a daily basis. After all the trades have been natched, the clearinghouse becomes the seller to all buyers and the buyer to all ellers, thereby guaranteeing the contractual obligations of each transaction. Thus, he clearinghouse provides security in transactions and absorbs settlement failures

should these occur (Hentschel and Smith 1994; Mengle 1995). The ability to absorb failures is dependent on the provisions for capitalizing the agents involved (Iben and Brotherton-Ratcliffe 1994). Exchanges can own their own clearinghouse, or the clearinghouse can be owned by other exchanges or financial institutions, such as banks

Exchange Regulation: Monitoring, Control, and Enforcement Functions

The proliferation of derivatives exchanges is a rather recent phenomenon, and governments are uncertain about how to regulate this market. Although there are signs of some tendencies toward convergence, the variations in legal and regulatory environments and in trade and business practices, together with the availability of derivatives products and the extent of domestic involvement in them, imply that systems will vary among countries. Despite the institutional choices available, the government's main objective is to monitor the exchange's activities and to set broad parameters for its operation. Minimum standards should be in place in areas such as contract design, market surveillance, reporting and record keeping, market transparency, safeguarding customer funds and assets, ensuring the financial integrity of the trading process, and protecting customers from fraud and markets from manipulation and trading abuses. Alongside these controls, a derivatives exchange needs to develop the capability to self-regulate by monitoring trading activities, ensuring contract execution, resolving disputes, enforcing rules and sanctions, and promoting professional conduct in order to increase investors' confidence (van der Bijl 1997).

Choice of Derivatives Products to Be Traded

To a large extent, the success of a derivatives exchange will depend on the choice of products to be traded The main categories of products are commodities, interest rates, currencies, individual stocks, and stock indexes The usual type of derivatives for these products are futures and options, although swap contracts are starting to appear in some exchanges The products to be traded need to have the following characteristics a sufficiently high level of price volatility to attract hedgers or speculators; a significant amount of money at risk, a significant number of domestic market participants—and possibly buyers and sellers from abroad, a large number of producers, processors, and banks interested in using derivatives contracts (that is, enough speculators to provide additional liquidity); and a weak correlation between the price of the underlying asset and the price of the already-traded derivatives contract(s) in other exchanges (basis risk). In the case of commodities, there also needs to be homogeneity in product specification, quality, and grade The creation of a single derivatives contract is conducive to increasing the contract's liquidity and hence its appeal to hedgers and speculators Note that financial products (that is, interest rates, currencies, and stocks) are almost by definition homogeneous products.

Results from an International Survey

This examination of derivatives exchanges is based on data from a survey that was distributed to 75 derivatives exchanges in 29 countries—almost all the exchanges that were in operation in 1996. Forty-two major exchanges in industrial and emerging markets answered the survey. Detailed results can be found in Tsetsekos and Varangis (1998), particularly in tables 1 through 11. Our analysis relied on the stage of development of the country's capital market to separate the respondents into emerging-market and industrial-country derivatives exchanges based on the International Finance Corporation's *Emerging Stock Market Factbook* (1988–95). Under this classification, 8 of the 42 respondents are considered to be emerging derivatives exchanges.

Chronology of Contract Introduction

The first derivatives contract was an agricultural contract introduced at the Chicago Board of Trade (CBOT) in 1859. That agricultural contracts were the first derivatives reflects that sector's importance in economic development during that period. Trading in derivatives of nonprecious metals began in 1878 at the London Metal Exchange. Globalization of transactions and exchange rate volatility led to the development of currency contracts in the late 1960s Increasing capital mobility and the development of capital markets in emerging economies promoted the development of interest rate contracts and contracts based on stocks and stock indexes. The Chicago Mercantile Exchange (CME) introduced currency derivatives in 1968 and interest rate derivatives in 1971. Shortly thereafter, the Chicago Board of Options Exchange offered equity-based derivatives. Energy indexes were introduced in 1974, stock indexes in 1978, and swaps in 1989.

Although exchanges had typically introduced derivative instruments for agricultural commodities first and more recently added interest rate, index-based, and equity derivatives products, emerging markets have typically introduced index-based and interest rate derivatives first. Index-based contracts are derivatives contracts whose underlying price is the value of an index of stocks. The U.S. indexes are the Dow Jones Industrial Average, the S&P 500, and the Nasdaq 100. Other major indexes include the Nikkei 225 index in Japan, the FTSE 100 in the United Kingdom, the DAX in Germany, and the CAC 40 in France. In terms both of total number of products traded globally in derivatives exchanges and of volumes traded, the most popular products are based on stock indexes, followed by interest-rate-based products, currencies, and, finally, agricultural commodities.

FINANCIAL DERIVATIVES PREFERRED TO COMMODITY DERIVATIVES. In view of the globalization of commodities markets and the liberalization of agricultural markets

in emerging economies, these countries' interest in establishing commodity derivatives may seem less justifiable. Several reasons, however, explain this apparent anomaly. First, the liberalization of commodities markets has increased the pass-through of international commodity prices to domestic commodities, making the use of derivatives contracts traded abroad feasible for domestic hedgers. In other words, the basis risk for using existing derivatives contracts has been declining. For example, Mexican wheat producers may now find that using the CBOT wheat futures contract is feasible for hedging Mexican wheat price risk. Thus, the need to develop a Mexican wheat contract may be reduced.

Second, financial markets are more country specific, and the demand for derivative instruments stems mainly from domestic users. Equity derivatives are based on equities in the local stock market; currency derivatives refer to the local currency relative to a foreign currency; and interest rate derivatives are based on the local bond market. More recently, institutional investors have taken an interest in these products. For example, to take a position in the local stock market, an institutional investor could purchase (or sell) futures on the equity index. If an investor wants to purchase a domestic stock but hedge that purchase against a drop that would affect stocks across the board, he could purchase a "put" option on the domestic equity index. Thus the market (country) specificity of financial instruments (stocks, bonds, and currencies) usually makes derivatives based on these instruments appear in local markets. The introduction of derivative instruments usually follows the development of domestic capital markets, but in some cases derivative instruments based on one market are traded on another. For example, the Nikkei 225 stock average is traded on the CME, while Italian and German bonds trade on the London International Financial Futures Exchange (LIFFE), and the Brazilian real trades on the CME.

Third, experience has shown that financial derivatives attract relatively higher liquidity than commodity derivatives. The Budapest Commodity Exchange experienced a dramatic increase in volume when it introduced financial derivatives contracts. In Brazil the São Paulo Commodities and Futures Exchange (BM&F) reported that the turnover of all agricultural contracts traded was about $3.5 billion for 1996, less than 1 percent of the total value of the $4.2 trillion contracts traded. At the CBOT and the CME, agricultural trades were less than 30 and 5 percent, respectively, of the total trading volumes of the exchanges.

TIMING OF SUCCESSIVE CONTRACT INTRODUCTION. According to survey respondents, about 36 months elapsed between the introduction of the first and second contracts, 18 months between the second and third, and 48 months between the third and fourth. The survey finds that introducing a derivatives contract takes relatively longer in emerging than in industrial markets. A noticeable difference is that emerging-market exchanges introduce (stock) index products relatively more quickly than do industrial markets.

The survey shows that agricultural derivatives products were the first to be delisted)y an exchange, followed by index products. The data do not cover exchange-traded)roducts that are inactive (that is, still listed but not traded, or traded but illiquid).

?xchange Structure

Most exchanges use an open outcry system, but electronic trading systems are in-.reasingly used. Twelve exchanges (of the 39 reporting) rely exclusively on electronic rading, and 8 others employ some form of electronic trading system. Recently es-ablished exchanges are more likely to use electronic-based systems for trading, an-icipating lower trading costs, which will be more attractive to businesses and inves-ors. An exchange using an electronic system could also draw business from traders round the world, significantly expanding the potential market. And advocates of lectronic trading say that it could be safer. For example, at the Beijing Commodity :xchange each trade is checked for adequate margins before the computer accepts it. \nd the BM&F in São Paulo performs back-office trade clearing and processing, a ask performed by member firms in most developed markets.

:LEARING ARRANGEMENTS. Most of the exchanges in the survey require initial mar-in deposits and margin calls (variation margins), with the exchange guaranteeing he contracts through its own clearinghouse. There is some uniformity in margin :quirements across the exchanges surveyed, and most require initial and variation ıargins with daily settlements. There are, however, disparities in the ways that mar-ın deposits are collected (whether gross or net) and in the collateral the clearing-ouse accepts—cash, securities, or letters of credit. Two-thirds of the exchanges sur-:yed own their own in-house clearing facilities; the remaining third cited ownership y banks, other financial institutions, or other exchanges.

In recent years exchanges have begun to explore a system of common clearing to nprove efficiency. In 1997 discussions picked up with an initiative by the Futures ıdustry Association to develop a proposal that meets the needs of all parties in-ɔlved. In Canada the Winnipeg Commodity Exchange is negotiating with the Ca-adian Derivatives Clearing Corp. (CDCC) to enter into a clearing services agree-ıent. CDCC clears for the Toronto and Montreal exchanges.

XCHANGE OWNERSHIP AND MEMBERSHIP. Most exchanges operate as nonprofit self-ɔverning entities owned by their members; a few operate as subsidiaries of other ɩchanges or as limited liability companies. Exchange membership distinguishes be-veen full (or regular) and associate (or limited) members. Only full members are ıtitled to voting privileges pertaining to self-regulation of the exchange. Associate ıembers have the same trading rights as full members, but often their activities do ɒt cover clearing transactions. Among the exchanges in the survey, the median

percentage of full or regular members is 80 percent, of which 76 percent have clearing rights. Only 16 percent of associate members have clearing rights.

Survey respondents indicated that the prevailing regulatory structure was exercised through a parliamentary law, under the authority of the government.

Economic and Capital Market Conditions

Using data from the International Monetary Fund's *International Financial Statistics* for various years, we examined the extent to which differences in economic and capital market conditions explain differences in the derivatives exchanges in industrial and emerging economies. As economic proxies, we used changes in consumer prices, prime interest rates, government bond yields, industrial production, growth in real gross national product (GNP), the level of GNP, and the share of investments in GNP. As proxies for capital market conditions, we used stock market turnover and capitalization, the variance in stock market capitalization, the value of stocks traded, the volatility in value traded, and the number of listed companies in the stock exchange, as reported by the International Finance Corporation in its *Emerging Stock Market Factbook.*

We found no statistically significant differences between emerging and industrial derivatives markets with respect to the economic proxies. We did find some differences in capital market conditions, but these may be explained by the size and maturity of industrial capital markets. We concluded that our tests give no conclusive indicators for the degree of market readiness for developing a derivatives exchange. In this respect, additional research is needed to establish whether such indicators exist and, if they do, to quantify them.

Launching New Derivatives Exchanges and Products

In considering the establishment of new derivatives exchanges, policymakers should begin with a feasibility analysis. Such an analysis should investigate the degree to which the preconditions discussed earlier are met, identify key areas of weakness, and formulate plans to address these areas. There is no point in proceeding with the design of the exchange and the products to be traded until this analysis is conducted The next step involves selecting the products to be traded and deciding on the elements of the exchange's microstructure. The research presented here, and experiences from various countries, point to some important lessons.

First, the chances of success are higher if index-based and interest-rate-based derivatives products are introduced first. Because our survey indicated that it takes more time to introduce derivatives contracts in emerging than in industrial markets, with the notable exception of derivatives based on an index of equities, emerging markets are likely to have more success in introducing financial derivatives related to equities.

Recent experience indicates that derivative instruments on agricultural products are more difficult to introduce because liberalization of agricultural markets tends to lag behind financial markets. Furthermore, financial markets tend to create much more liquidity than agricultural markets. For instance, in 1996 the Budapest Commodity Exchange increased its liquidity by 400 percent when it introduced financial contracts (mainly currency) to its product line. And because of the globalization of commodity markets, the potential for using existing contracts in established exchanges reduces the need to develop local agricultural contracts. As a general rule, emerging-market economies that have relatively successful commodity exchanges have sizable local commodity markets (for example, Argentina, Brazil, China, and Malaysia).

Second, appropriate regulations and a conducive legal environment are crucial for the development of derivatives exchanges. The literature attributes problems in the legal-regulatory infrastructure as major impediments in the drive to develop derivatives exchanges in emerging markets. The most important of these problems are:

- Antagonism between market sectors (banking, derivatives, and securities) over which entity should regulate and supervise the exchange—and under what rules
- A lack of confidence as a result of scandals, corruption, and market failures
- Uncertainty about the equitable application of laws and regulations, the enforceability of obligations, and the lack of market-oriented insolvency laws.

Third, partnerships and joint ventures between new and existing exchanges can be mutually beneficial. The CBOT is considering such ventures with exchanges in Argentina, Poland, and Turkey. Established exchanges can offer technology and know-how, and their members in turn can gain access to a potentially high-growth market. In fact, rather than setting up their own derivatives exchanges, several emerging economies could do better by using other well-established exchanges and listing their products with them. Regional exchanges such as the Stockholm-based Options Market offer another way to improve market liquidity, but they may be harder to develop and coordinate.

Fourth, policymakers in emerging economies should look at whether access provided to market participants in their countries allows them to trade in derivatives exchanges abroad. Even where there are local exchanges, removing the barriers to overseas trading could increase the liquidity of the local exchange by providing opportunities for arbitrage between local and foreign exchanges.

Fifth, electronic trading appears to be the choice among new exchanges. Lower transaction costs for users are often cited as the key benefit of electronic trading.

To some extent, derivatives markets complement developments in the stock markets. By the end of 1996, more than 78 developing countries had stock markets; during the 1990s their capitalization increased more than 10 times, and the number of domestic companies listed more than doubled. Derivatives exchanges have played a major role in these developments. They have contributed to more balanced alloca-

tion of resources and have enabled the transfer of risk within a country and even across countries. Although there are concerns about the explosive growth of derivatives and the risks that they may create, business at these exchanges is increasing, and the exchanges continue to grow and to introduce new products.

Appendix. Types of Derivative Instruments

Forward Contracts

A forward contract is an agreement to purchase or sell a given asset at a future date at a fixed (predetermined) price. In a forward contract the buyer and the seller assume each other's performance risk. The terms and conditions of the forward contract are usually specific to each transaction, although some forward contracts are standardized.

Futures Contracts

A futures contract is similar to a forward contract: the buyer (seller) of a futures contract agrees to purchase (sell) a specified amount of an asset at a specified price on a specified date. But futures contracts differ significantly from forward contracts in several ways. First, contract terms (amounts, grades, delivery dates, and so forth) are generally standardized. Second, transactions are handled only by organized exchanges through a clearinghouse system. Third, profits and losses in trades are settled daily (marked to market). Fourth, futures contracts require depositing a certain amount of margin money in the exchange as collateral. Fifth, while forward contracts involve delivery (exchange) at maturity, futures contracts are usually closed before that. Thus, futures contracts separate the purchase and sale of assets from hedging. Through these arrangements, futures contracts significantly reduce the credit and default risk entailed in forward transactions. Contract standardization also improves liquidity (that is, the contract volumes traded).

Options Contracts

An option on a futures contract is the right—but not the obligation—to purchase or sell a specified quantity of the underlying futures contract at a predetermined (strike) price on or before a given date. Exchange-traded options, like futures contracts, are standardized. There are also over-the-counter options offered by banks and brokers, which can be customized. The purchase of an option is equivalent to price insurance; therefore, there is a price to be paid (just like an insurance premium). Some important definitions regarding options are:

Call. A call option gives the buyer the right, but not the obligation, to buy the underlying futures contract at a predetermined price during a given period of time. Call options are usually purchased as insurance against price increases.

Put. A put option gives the buyer the right, but not the obligation, to sell the underlying futures contract at a predetermined price during a given period of time. Put options are usually purchased as insurance against price declines.

Strike or exercise price. The price at which the futures contract underlying an option can be purchased (if a call) or sold (if a put).

Premium. The price paid for the options contract.

Exercise. To exercise a call (or put) is to buy (or sell) the underlying futures contract at the strike price.

Time to expiration. An option is good only for the length of time specified in the contract. The last day that an option can be exercised is called the expiration date.

Swap Contracts

A swap contract is an agreement to exchange, or swap, a floating price or rate for a fixed price or rate (or vice versa) for an asset at specific time intervals. A swap is like a series of forward contracts lined up on a schedule, but it does not involve physical exchange of assets. Swaps solve problems relating to the need for longer-term price fixation, but they tend to be credit-intensive and carry the risk of nonperformance.

Notes

George Tsetsekos is vice-provost and professor of finance at Drexel University, and Panos Varangis is an economist with the Development Research Group of the World Bank Earlier versions of the paper were presented at the Annual Meetings of the Multinational Finance Society, June 25–28, 1997, and at the Annual Meetings of the Financial Management Association, Honolulu, Hawaii, October 15–18, 1997 The paper is based on a World Bank–sponsored research project (RPO 680-45) entitled "Market Architecture and Design of Derivatives Exchanges " The authors acknowledge the assistance of Tom Scott, Stan B Gyoshev, and Yiannis N Koulouriotis in completing this research They would also like to thank Richard DeFusco and P C Kumar for their valuable comments on earlier drafts of the paper.

1. According to the Bank for International Settlements (BIS), at the end of 1996 the notional principal outstanding of financial derivatives approached $35 trillion, of which approximately $10 trillion was in exchange-traded instruments, the rest was traded over the counter

References

Abken, Peter A 1994. "Over-the-Counter Financial Derivatives Risky Business?" *Economic Review* 79(2).1–22. Federal Reserve Bank of Atlanta

Becketti, Sean. 1993 "Are Derivatives Too Risky for Banks?" *Economic Review* 78(3):27–42. Feder, Reserve Bank of Kansas City.

Chang, H , Laurie Kaplan, and Paul Knapp. 1999. "Commodity Futures Exchanges in Developin Economies: Conditions for Success." Catalyst Institute, Chicago, Ill. Processed.

Chapman, Alger B. 1994 "Future of the Derivatives Markets. Products, Technology, and Partici pants " *Financial Practice and Education* 4(2)·124–28

Domowitz, Ian. 1995 "Electronic Derivatives Exchanges: Implicit Mergers, Network Externalitie and Standardization." *Quarterly Review of Economics and Finance* 35(2):163–75.

Garber, Peter M 1998 *Derivatives in International Capital Flows.* Working Paper 6623. Cambridge Mass · National Bureau of Economic Research.

Gennotte, Gerard, and Hayne Leland. 1994 "Low Margins, Derivative Securities, and Volatility *Review of Futures Markets* 13(3).709–42.

Hentschel, Ludger, and Clifford W Smith. 1994. "Risk and Regulation in Derivatives Markets *Journal of Applied Corporate Finance* 7(3):8–21.

Iben, Ben, and Rupert Brotherton-Ratcliffe 1994 "Credit Loss Distributions and Required Capita for Derivatives Portfolios " *Journal of Fixed Income* 4(1):6–14

Leuthold, R M 1992 "Access to Futures and Options Markets Own versus Existing Markets Paper prepared for the Conference on Risk Management in Liberalizing Economies Issue of Access to Futures and Options Markets, Association Française des Banques, Paris, June 3–5 Processed

Mengle, David L. 1995 "Regulatory Solutions to Payment System Risks. Lessons from Privatel Negotiated Derivatives " *Journal of Financial Services Research* 9(3)·381–91

Merton, Robert C , and Zvi Bodie. 1995 *The Global Financial System A Functional Perspective* Boston, Mass Harvard Business School Press

Peck, A E 1985. "The Economic Role of Traditional Commodity Futures Markets " In A E Peck ed , *Futures Markets Their Economic Role* Washington, D C American Enterprise Institute fo Public Policy Research

Pirrong, S C 1994 "Derivatives Exchanges, Liquidity and Locals A Look to the Future " Catalys Institute, Chicago, Ill. Processed

Remolona, Eli M. 1992 "The Recent Growth of Financial Derivative Markets " *Quarterly Review* 17(4) 28–43. Federal Reserve Bank of New York

Scott, Louis O 1992 "The Information Content of Prices in Derivative Security Markets " *IMF Staff Papers* 39(3).596–625

Smith, Clifford W , Jr , Charles W Smithson, and D Sykes Wilford 1990 *Managing Financia Risk* New York Harper & Row

Stout Lynn A 1996 "Derivatives Trading in a World of Risk and Uncertainty " *Brookings Review* 14(1) 38–41

Tsetsekos, George, and Panos Varangis 1998 "The Structure of Derivatives Exchanges Lesson from Developed and Emerging Markets " Policy Research Working Paper No 1887 World Bank Development Economics Department, Washington, D C

van der Bijl, R 1997 "Exchange-Traded Derivatives in Emerging Markets An Overview " Presenta tion at the Commodity Futures Trading Commission International Regulatory Training Semi nar, Chicago, Ill , October 20 Processed

Growth and Institutions: A Review of the Evidence

Janine Aron

Africa's disappointing economic performance, the East Asian financial crisis, and the weak record of the former Soviet Union have focused attention on the role of institutions in determining a country's economic growth This article critically reviews the literature that tries to link quantitative measures of institutions, such as civil liberties and property rights, with growth of gross domestic product across countries and over time. An important distinction is made between indicators that measure the performance or quality of institutions and those that measure political and social characteristics and political instability. The evidence suggests a link between the quality of institutions and investment and growth, but the evidence is by no means robust.

> *I wish to assert a much more fundamental role for institutions in societies; they are the underlying determinant of the long-run performance of economies.*
>
> North 1990: 107

The role of institutions in promoting growth in developing and emerging economies has sparked renewed interest in recent years (World Bank 1993, 1997; Stiglitz 1998). A burgeoning literature thus seeks to determine the extent to which the quality of public and private economic institutions, the particular structure of governance, and the extent of social capital (or civic engagement) affect growth. Evidence from global cross-country econometric studies is potentially important because the paucity and weakness of both macroeconomic and institutional data for many developing countries preclude robust policy interpretations on a country-by-country basis (Srinivasan 1995; Lal and Myint 1996).

If there is clear evidence that weak political and economic institutions significantly hamper growth, policymakers might propose measures that strengthen institutions in particular ways or that encourage more appropriate political structures (Aron 1996). For instance, African countries usually form a significant part of the

The World Bank Research Observer, vol 15, no 1 (February 2000), pp 99–135

cross-country samples, and the region plays a distinctive role in this literature. Many African countries possess very weak public and private institutional frameworks, and Sub-Saharan Africa has experienced the slowest economic growth of any region in the world, with poverty large and deepening.[1] Several cross-country studies of growth (for example, Easterly and Levine 1997) have found that the conventional factors of growth (labor, physical and human capital accumulation, and so on) do not fully explain Africa's experience and have turned to an institutional explanation.

Unfortunately, interpreting the evidence on growth using institutional measures is not a straightforward matter. First, the process of integrating institutions and institutional change into economic theory is of comparatively recent vintage. The growth literature does not subscribe to one overarching definition of economic, political, and social institutions, their processes of change, and their likely channels of influence on economic outcomes. Yet, as North (1990: vii) suggests, "The specification of exactly what institutions are, how they differ from organizations, and how they influence transaction and production costs is the key to much of the analysis." In the empirical literature the terms politics and institutions encompass a wide range of indicators, including institutional quality (the enforcement of property rights), political instability (riots, coups, civil wars), characteristics of political regimes (elections, constitutions, executive powers), social capital (the extent of civic activity and organizations), and social characteristics (differences in income and in ethnic, religious, and historical background).[2] Economists often rely on several of these types of indicators to capture the features of institutions, although each has a potentially different channel of impact on growth.

Second, the growth literature is burdened by a range of serious problems with data, methodology, and identification, which many authors underestimate or choose to ignore (see the survey in Temple 1999). As Heston and Summers (1996) note, there has been an astonishing proliferation of research based on the Penn World Tables data. Yet, although various studies establish statistically significant relationships, they frequently do not test the sensitivity of their results to different model specifications, data outliers, measurement errors, reverse causality between regressors and growth, and bias due to the possible omission of variables.

Finally, much of the cross-country empirical literature on growth, while loosely related to the structural growth models of Solow (1956) and Mankiw, Romer, and Weil (1992), in fact consists of reduced-form growth regressions (explained in box 1). The structural Solow growth equation (explained in box 2) includes current investment as a determinant of growth, as well as a range of other variables (such as population growth, income, and human capital). Where the variables that measure the quality of institutions enter this equation, they can be interpreted as having a direct effect on growth because they improve the efficiency of investment. Since investment is already included in the growth equation, any effects that institutional variables exert on growth through an increase in the volume of investment are indi-

Box 1. *Structural Growth Equations versus Reduced-Form Growth Equations*

In the literature on growth and institutions, it is important to distinguish between the results from *structural* models and those from *reduced-form* models

The complete structural system for a growth model consists of at least two equations, usually derived from economic theory—typically, equations for current growth and investment Although current investment is a determinant of current growth, there also is reverse causality For instance, high current growth may have a demand effect on investment and may also increase saving, which in turn helps to finance investment. The variables whose value is determined within the system—namely, growth and investment—are called *endogenous* variables, those outside the system, on which the system has no influence, are called *exogenous* variables For example, population usually is treated as *exogenous*. Examples of *lagged* endogenous variables in growth models are beginning-of-period values for income or the level of human or physical capital

Single structural equations Many empirical growth models are based on single structural growth equations (for example, the Solow growth equation). Since there is reverse causality between current growth and investment, it becomes important to correct for bias (or overestimation) in calculating the importance of investment for growth Most single-equation growth studies fail to make this correction A discussion of these issues in some of the literature on politics and institutions appears in Deaton and Miller (1996).

Reduced-form models In reduced-form models, the endogenous dependent variable is expressed only in terms of exogenous variables and parameters For example, endogenous investment on the right-hand side of the equation is replaced by a set of variables fully determining investment Thus the reduced form describes the results of interactions among the endogenous variables but not the simultaneous interactions themselves The important disadvantage of these models, generally speaking, is that one loses the ability to distinguish the different channels of influence on growth (for example, to quantify the effect of investment on growth). Moreover, in many single-equation reduced-form growth equations in the literature, the set of variables determining investment is incomplete or contains other current endogenous variables

rect. A separate investment equation would be necessary to ascertain the direct effects of institutional variables on the volume of investment.

However, no problem arises in connection with the simultaneity of investment and growth when reduced-form growth regressions are used because these regressions omit the investment variable and replace it with the set of variables that determine investment. For sufficiently general models, this approach should capture both the influences on the volume of investment and the efficiency of investment (that is, both the direct and indirect influences on growth). Unfortunately, therefore, when institutional quality variables are included, the different effects of institutions on growth are conflated and difficult to disentangle. Furthermore, in practice these models are rarely sufficiently general.

This review examines a range of influential studies in the heterogeneous literature on growth and institutions, both to obtain a better understanding of the linkages involved and to assess critically the sometimes strong claims made by the authors. It argues that a sensible interpretation of the effects of politics and institutions on eco-

Box 2. *How Do Institutions and Politics Enter the Solow Growth Model?*

Solow's model of economic growth is based on the premise that output in an economy is produced using a combination of labor (L) and capital (K), under constant returns, so that doubling inputs results in a doubling of output. Modern versions of the story distinguish between physical and human capital. The quantity of output (y) is also determined by the efficiency (A) with which capital and labor are used, or

$$y = Af(L,K). \quad (1)$$

With the further assumption of competitive markets, the growth rate of the economy, then, is a weighted sum of the growth rates of the efficiency parameter, g_A (sometimes called technical progress) of the labor force, g_L, and of the capital stock, g_K, where the weights on the latter two are the shares of payments to labor and capital in gross domestic product:

$$g_y = g_A + a_L g_L + a_K g_K \quad (2)$$

A critical assumption in the Solow model is that the marginal product of capital decreases with the amount of capital in the economy. In the long run, as the economy accumulates more and more capital, g_K approaches 0, and the growth rate is determined by technical progress and growth in the labor force. But in the short run, an economy that accumulates capital faster will enjoy a higher level of output.

Cross-country growth regressions study the determinants of growth by using each country's experience as a data point. Typically, there are two types of regressions: (a) those (the majority) that estimate changes in levels of output—essentially estimating equation (1) at two points in time—and (b) those that estimate equation (2), linking growth rates over several decades with technical progress (assumed constant across countries) and with labor force and capital stock growth rates

How can country institutional differences enter into these regressions? For both categories, the quality of institutions can affect technical progress For instance, David (1997) discusses how take-up of technology is constrained by "social capability" Thus the rate of technical progress is no longer constant across countries, rather, it depends on country-specific institutional differences. For those in the first category, the country's initial level of technical efficiency may be affected by the quality of its institutions. This in turn will affect the efficiency of investment. Regressions that neglect the role of initial technical efficiency (due to institutions) may then overstate the role of investment in economic growth Finally, in many developing countries there are threshold levels of infrastructure, property rights, and education (all of which are sensitive to institutions) that must be met before production is feasible. The presence of these thresholds means that the constant-returns assumption may not hold. Thus, introducing country-specific institutional variables can affect the estimate of the responsiveness of output to capital (human as well as physical) in both categories of regressions.

nomic growth requires a more structural theoretical approach based on a common institutional framework.

What Is Meant by the Term "Institutions," and How Are They Measured?

Proxies for institutions were first introduced into cross-country growth and investment equations more than a decade ago, and recently this literature has experienced

a renaissance. Researchers have used diverse measures, encompassing political instability, the attributes of political institutions, social characteristics and social capital, and measures of the quality of institutions that affect economic exchange. The literature on economic growth typically has classified and treated these proxies collectively as "sociopolitical measures." This practice has tended to obscure the different channels through which institutions operate and has impoverished the interpretation of the role of institutions in growth. This is a serious flaw in analyzing developing countries, where weak institutions are implicated in low growth.

Defining Institutions

Classifying schools of thought in the relatively recent tradition of political economics, Posen (1998) indicates that the considerable influence of North (1990) and his school largely has affected political scientists and economic historians, not economists as such. Yet North provides a clear institutional framework that may be integrated with the burgeoning economic literature to provide a richer interpretation of the effect of institutions on growth. Consider the following observation:

> We have only to contrast the organisation of production in a Third World economy with that of an advanced industrial economy to be impressed by the consequences of poorly defined and/or ineffective property rights. Not only will the institutional framework result in high costs of transacting in the former, but also insecure property rights will result in using technologies that employ little capital and do not entail long-term agreements. . Moreover, such mundane problems as the inability to get spare parts or a two-year wait to get a telephone installed will necessitate a different organisation of production than an advanced country requires. A bribe sufficient to get quick delivery through the maze of import controls or get rapid telephone installation may exist, but the resultant shadow transactions costs significantly alter relative prices and consequently the technology employed. (North 1990: 65)

The institutional framework comprises both formal and informal constraints. North describes a continuum with unwritten taboos, customs, and traditions at one end and constitutions and laws governing economics and politics at the other. In the absence of formal rules, a dense social network leads to the development of customs, laws, trust, and normative rules that constitute an informal institutional framework (for example, see Bates 1989). Naturally, informal constraints on behavior are pervasive and important in modern economies too (David 1994, 1997).

With economic development comes a unidirectional move along the continuum as increasing specialization and the division of labor associated with more complex societies raise the rate of return to formalizing political, judicial, and economic rules

and contracts that facilitate political or economic exchanges. There is a hierarchy of such rules: from constitutions, to state and common laws, to specific by-laws, to individual contracts. Usually, the higher the rules lie in the hierarchy, the more costly they are to alter.

WHY ARE INSTITUTIONS WEAKER IN SOME COUNTRIES THAN IN OTHERS? Institutions may be weak because rules simply are absent, rules are suboptimal, or useful rules are poorly enforced. North emphasizes the cost of information, since resources are required not only to measure the attributes of a good or service in economic exchange but also to define and measure the rights that are transferred and to protect these rights by policing and enforcing agreements. The more complex is the exchange, the more costly are the institutions. Where such costs are prohibitive (given the technology), rules may not be worth devising, and ownership rights remain undefined. Of course, changes in technology or in relative prices may alter the relative gains from devising rules. Where rules do exist, they may be counterproductive (excessive import controls, for instance). Finally, useful rules may not be enforced when the costs of monitoring and enforcement prove too high. For instance, constitutions abound in Africa, but many are ineffective. Similarly, a lack of correlation between the central bank's constitutional autonomy and low inflation in developing countries is likely due to weakness of the judiciary in enforcing autonomy (Cukierman, Webb, and Neyapti 1992).

HOW IS ECONOMIC GROWTH AFFECTED BY THE INSTITUTIONAL FRAMEWORK? The structure of both formal and informal rules and the character of their enforcement are what define the incentives and wealth-maximizing opportunities of individuals and organizations. North (1990: 110) asserts, "Third World countries are poor because the institutional constraints define a set of payoffs to political/economic activity that do not encourage productive activity." Such rules affect both individuals and organizations, defined as political organizations (city councils, regulatory agencies, political parties, tribal councils), economic organizations (firms, trade unions, family farms, cooperatives, rotating credit groups), educational bodies (schools, universities, vocational training centers), and social organizations (churches, clubs, civic associations).

The institutional framework affects growth because it is integral to the amount spent on both the costs of transactions and the costs of transformation (in the production process). Transaction costs, for example, are far higher when property rights or the rule of law are not reliable. In such situations private firms typically operate on a small scale, perhaps illegally in an underground economy, and may rely on bribery and corruption to facilitate operations. Transformation costs, too, can be raised substantially because unenforceable contracts mean using inexpensive technology and operating less efficiently and competitively on a short-term horizon.

When institutions are poorly defined or there are few formal institutions, economic activities are restricted to interpersonal exchanges. In such cases, repeat activities and cultural homogeneity facilitate self-enforcement. Transaction costs may be low in such an environment, but transformation costs are high because the economy operates at a very low level of specialization. Economic exchange also could operate at one remove, via social networks, but contracts still are constrained by kinship ties. It is clear, however, that firms or agents in an environment of weak institutions cannot engage in complex, long-term, and multiple-contract exchanges with effective enforcement as they do in industrial economies. A basic structure of property rights that encourages long-term contracting appears essential for the creation of capital markets and economic growth.

Rules with the wrong incentives can be hard to change when the costs are too high. There may be considerable sunk costs in developing political and economic organizations that can operate in a weak institutional environment. It is conceivable that such organizations ultimately could force an improvement in institutions, but unless a certain threshold of inefficiency is reached, they are more likely to perpetuate and entrench the weak institutional environment. Any organizations that develop in this framework may then become more efficient, but more efficient at making the society even less productive.

Classifying Institutions by Empirical Measures

The political science literature has been constructing and testing measures of political instability, political and civil freedom, and the characteristics of democracy for many years (Przeworski and Limongi 1993). Kormendi and Meguire (1985) and Scully (1988) are among the first papers to explore the effect of variables describing qualitative political and civil liberties on cross-country growth and investment. More recently, the literature has adapted a number of the original indexes but is increasingly using disaggregated measures by country risk-rating or credit-rating agencies using subjective surveys to capture more focused economic notions of institutional quality, such as respect for property rights.

The way in which the empirical institutional measures are categorized is important for interpreting their effects. The theory-based classification adopted here differs from related empirical surveys by Alesina and Perotti (1994) and Brunetti (1997) in that it distinguishes between measures that describe the attributes of institutions and those that evaluate their performance.[3]

Descriptions of the features of political and economic institutions—such as the presence or absence of constitutional rights—say nothing about how well such institutions perform. In contrast, measures of the quality of formal and informal institutions indicate how effectively the existing institutional rules or norms are implemented. For example, measures of the quality of formal institutions include

subjective rankings of the effectiveness of property rights and of the bureaucracy (that is, the ease of doing business), often drawn from cross-country surveys conducted by risk agencies. Such measures are proxies for the transaction and transformation costs of production that may affect the volume and efficiency of investment and hence growth (there also may be reverse causality from growth and investment to institutions).

It is difficult to measure informal constraints, but Putnam (1993) provides some measures of social capital that capture the extent (or perhaps the quality) of civic activity and organization. As indicated by North, informal social constraints based on the trust and norms inherent in networks and associations may influence growth both directly and indirectly (as with formal constraints). Some measures of social capital also reflect the ability of citizens to hold the state accountable. This highlights the role of external domestic and international agencies in promoting good governance, including civil society (Collier 1996). An example of such a measure is the subjective Gastil index of civil freedoms (including freedom of the press and of assembly).[4] Again, there could be a reverse causality from growth and investment to social capital.

It also is important to distinguish between the performance of formal and informal institutions and various measures of social characteristics and political instability. Clearly, although extremes of political instability leading to riots and civil war can destroy existing investment, the way in which instability affects the volume and efficiency of new investment is probably through limits on the effectiveness of formal and informal institutions. Social measures such as the degree of ethnic diversity in a country may well explain the tendency to experience political instability (or the duration of different political regimes) or could have a direct impact on the character of social capital and of formal institutions.

Thus social characteristics are perhaps more appropriate in explaining the duration of political regimes, while political instability measures and social characteristics may explain formal and informal property rights. Note again the possibilities for reverse causality, although social characteristics are fairly exogenous measures (Aron 1998).

Table 1 classifies the components of indicators of institutional measures in five categories: quality of formal institutions (typically drawn from surveys or risk ratings by investors); measures of social capital, which capture the intensity of social participation and organization; measures of social characteristics, including ethnic, cultural, historical, and religious categories; characteristics of political institutions, including constitutional rights and descriptions of the type of regime (dictatorship, democracy); and measures of political instability, including riots, strikes, civil war, duration of regime, and changes in the executive. The table also differentiates subjective measures, based on surveys and personal assessments, and objective measures,

(Text continues on page 114)

Table 1. *Political, Economic, and Social Measures from the Empirical Growth Literature*

Institutional measure	*Source*	*Period, country*	*Components of index*[a]	*Growth references using the measures*
A Institutional quality measures				
BERI disaggregated business risk indicators subjective (ranked by a "permanent" panel of experts)	Knack and Keefer (1995), data from Business Environmental Risk Intelligence (BERI) private firm for potential foreign investors	Annual from 1972, about 47 countries (7 African countries), not all countries start in 1972	*Security of contract and property rights* Bureaucratic delay, nationalization potential, contract enforceability, infrastructure quality	Knack and Keefer (1995, 1997), Barro (1996a); Clague and others (1996); Hassan and Sarna (1996), Knack (1996); Lane and Tornell (1996)
ICRG disaggregated business risk indicators subjective (ranked by staff of political risk services)	Knack and Keefer (1995), data from International Country Risk Guide (ICRG) private firm for potential foreign investors	Annual from 1982, 135 countries (34 African countries), not all countries start in 1982	*Security of contract and property rights* Rule of law, corruption in government, quality of the bureaucracy, repudiation of contracts by government, expropriation risk of private investment	Knack and Keefer (1995, 1977), Barro (1996a), Clague and others (1996), Hassan and Sarna (1996), Knack (1996), Lane and Tornell (1996), Sachs and Warner (1997)
Business International disaggregated risk indicators subjective (ranked by local observers)	Mauro (1995), data from Business International (BI) private firm for potential foreign investors, now incorporated into the Economist Intelligence Unit	1971–79, annual 57 countries 1980–83, annual, 68 countries (10 African countries)	*Institutional quality* Corruption index, bureaucratic efficiency sum of three measures (efficiency of judicial system, absence of red tape and absence of corruption), political stability sum of six measures (institutional change, social change, opposition takeover, stability of labor, relationship with neighboring countries, terrorism), institutional efficiency sums all nine	Mauro (1995), Clague and others (1996), Helliwell (1996a)

(Table continues on the following pages)

Table 1 *(continued)*

Institutional measure	*Source*	*Period, country*	*Components of index*[a]	*Growth references using the measures*
Borner, Brunetti, and Weder's political credibility index subjective (ranked by local entrepreneurs)	Borner, Brunetti, and Weder (1995), based on own 1992–93 survey of entrepreneurs in 28 countries	1981–90, 28 countries (8 African countries)	*Political credibility* Unexpected changes in laws and policies regularly affecting business, expectation that government sticks to major announced policies, changes in uncertainties over lawmaking in the last decade	Borner, Brunetti, and Weder (1995), Brunetti, Kisunko, and Weder (1998)
Measure of contract-intensive money objective	Clague and others (1995)	Uses available data in *International Financial Statistics* (IMF)	*Poor enforcement of contracts or property rights* M2 – C/M2, M2 is broad money, and C is currency outside banks (increases with efficiency)	Clague and others (1995, 1996)
Heritage Foundation index of economic freedom partly subjective (but not from risk-rating agencies)	Heritage Foundation, Washington, D C	Annual, from 1996, 161 countries	*Dimensions of market efficiency* Trade policy; taxation, government intervention, monetary policy, capital flows and foreign investment regulations, banking regulations; wage or price controls, protection of property rights, efficiency of regulation, extent of parallel market	Ng and Yeats (1999)
Gastil's political rights index (Freedom House index)· subjective (ranked, but not by local observers)	Gastil (1989, 1991), based on published and unpublished information about individual countries	Annual, from 1973, 165 countries	*Political rights measure (sometimes called "democracy")* Meaningful election of chief authority; meaningful election of legislature, fair campaigning; fair reflection of voter preference, multiple political parties, no military control, decentralized political power, informal consensus, significant opposition vote, recent shift in power through elections, no denial of self-determination of major groups	Kormendi and Meguire (1985), Scully (1988); Barro and Lee (1994), Helliwell (1994); Sachs and Warner (1995), Savvides (1995), Alesina and others (1996), Barro (1996a), Ghura and Hadjimichael (1996), Perotti (1996), Isham Kaufmann and Pritchett (1997)

Euromoney country risk disaggregated ratings mixed objective (financial market conditions) and subjective (ranked by average of evaluations of international panel of 35 bankers and analysts)	*Euromoney* (various issues)	Annual from 1970 (35 African countries) Disaggregated data from 1992 (changed methods of calculation of index in 1987)	*Components of country risk* Economic performance projections, political risk (subjective), debt indicators, debt in default or rescheduled; average of Moody or Standard and Poor credit ratings (part subjective), access to bank lending, access to short-term finance; access to capital markets (part subjective), discount on forfeiting *Property rights proxy ("political risk")* Risk of nonpayment or nonservicing of payment of goods, services, loans, trade-related finance and dividends and of nonrepatriation of capital	
Institutional Investor rating of risk of default on sovereign debt subjective (ranked by international panel of bankers)	*Institutional Investor* (semiannual publication)	March and September, from 1979, more than 100 countries (25 African countries)	*Property and contracts rights* Credit rating	Clague and others (1996)
B Social capital measures				
Gastil's civil liberties index (Freedom House index) subjective (ranked, but not by local observers)	Gastil (1989, 1991) based on published and unpublished information about individual countries	Annual, from 1973, 165 countries	*Civil liberties measure* No censorship, open public discussion, freedom of assembly; freedom of political organization, guaranteed socioeconomic rights, no gross inequality; no government indifference, no political terrorism, rule of law in political cases, freedom of religion, free trade unions, freedom of business, guaranteed personal rights; freedom of private organizations	Scully (1988), Levine and Renelt (1992), Barro and Lee (1994), Helliwell (1994), Sachs and Warner (1995), Savvides (1995), Alesina and others (1996), Barro (1996a), Ghura and Hadjimichael (1996), Perotti (1996), Isham, Kaufmann, and Pritchett (1997)

(Table continues on the following pages)

Table 1 *(continued)*

Institutional measure	*Source*	*Period, country*	*Components of index*[a]	*Growth references using the measures*
Gastil-Wright economic freedom ratings subjective (ranked, but not by local observers)	Wright (1982)	One year only, 165 countries	*Economic freedom* Freedom of property, freedom of association, freedom of movement, freedom of information	Spindler (1991)
Banks's government purges and peaceful demonstrations objective	Banks (1975 onward)	1970 onward, currently more than 190 countries	*Social capital measures* Purges any systematic elimination by jailing or execution of political opposition within the ranks of the regime or the opposition Antigovernment demonstrations any peaceful public gathering of at least 100 people for the primary purpose of displaying or voicing their opposition to government policies or authority, excluding demonstrations of a distinctly antiforeign nature	Easterly and Levine (1997)
World Values Survey subjective (surveys)	Inglehart (1994a), indexes in Knack and Keefer (1997)	1981, 1990–91 surveys of countries, more than 43	*Cultural values* Thrift, determination, hard work, obedience, religious faith, respect *Social capital measures* Associational relationships, trust	Inglehart (1994b), Helliwell (1996a, 1996b), Knack and Keefer (1997)
Putnam's social capital and local government and institutional performance measures mixed objective and subjective (local surveys)	Putnam (1993)	1860–1987, Italian regions	*Civic community* Preference voting, referendum turnout, newspaper readership, number of sports or cultural associations *Institutional performance* Reform legislation, day care centers, housing and urban development, statistical and information services, legislative innovation, cabinet stability, family clinics, bureaucratic responsiveness industrial policy instruments budget promptness local health unit	Helliwell and Putnam (1995)

			spending, agricultural spending *Local government performance* Implementation of 15 communal facilities including sports, sewers, libraries, trash collection, water system, school transport, day care centers	
C Social characteristics				
Taylor and Hudson's ethnolinguistic fractionalization measure objective	Mauro (1995), based on data from Taylor and Hudson (1972)	1960, 113 countries (12 African countries)	*Ethnolinguistic fractionalization* Probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group	Many studies, including Mauro (1995), Perotti (1996); Easterly and Levine (1997), Collier (1999)
Easterly and Levine's ethnic diversity and tension measures objective	Easterly and Levine (1997) data set, from various sources	Specific years as indicated, more than 75 countries	*Measures of ethnic tension* Racial tension in 1984, proportion of population belonging to minorities at risk in 1990, percentage of population not speaking country's official language at home; percentage of population belonging to separatist movements in 1960 and 1975	Easterly and Levine (1997)
Adelman-Morris index of social development objective, plus subjective exclusion of two indicators	Adelman and Morris (1968)	1957–62, 73 countries (26 African countries)	*Social development and capability* Size of traditional agricultural sector, extent of dualism, extent of urbanization; character of basic social organization, importance of indigenous middle class, extent of social mobility, extent of literacy; extent of mass communication, crude fertility rate, degree of modernization of outlook	Temple and Johnson (1998)
Barro's measures of colonial status and religious affiliation, objective	Barro (1996a)	138 and 136 countries, respectively	*Social requisites for democracy* Colonial status based on most recent ruler after 1776, religious affiliation by majority religion in a seven-way breakdown (proportion of population)	Barro (1996a)

(Table continues on the following pages.)

Table 1 *(continued)*

Institutional measure	*Source*	*Period, country*	*Components of index*[a]	*Growth references using the measures*
D Political characteristics				
Clague, Keefer, Knack and Olsen's political regime indicators objective	Clague and others (1996), constructed from Banks (1979), Gurr (1990) to 1986 and extrapolated to 1990 with *Europa Yearbook*	Around 1969–90	*Type and duration of political regime* Dictatorship, almost dictatorship, intermediate category, almost democracy, and democracy; type of regime based on rankings from summing outcomes from Gurr's and Banks's measures of executive competitiveness, selection, and legislative effectiveness, duration variables refer to numbers of consecutive years spent in regimes, resetting variables when status changes	Clague and others (1996)
De Vanssay and Spindler's constitutional rights indicators objective	De Vanssay and Spindler (1992)	OECD plus non-OECD countries	*19 constitutional variables* Bill of rights, right to privacy, right to unionize; political attributes, such as whether supreme court has final constitutional authority	De Vanssay and Spindler (1992)
Bates and others' measures of political transition objective	Bates and others (1996), work in progress	1970–91, 49 countries (all African)	*Measures of political transition* Executive scale; legislative scale (further scales in progress)	None as yet
Bollen's democracy measure objective	Bollen (1990), drawing on Banks (1979) and Taylor and Hudson (1972)	1960, 1965, more than 110 countries	*Political components ("democracy")* Three concerning political liberties, three concerning political rights	Many studies, including Helliwell (1994), Barro (1996a)
E Political instability				
Taylor and Jodice's and Banks's political instability indicators objective	Persson and Tabellini (1994), using Taylor and Jodice (1983, 1988) and Banks (various issues)	Annual, from 1960, 136 countries	*Political instability and characteristics* Number of revolutions, successful coups, unsuccessful coups, and political assassinations, number of changes in the composition of the executive, number of riots and demonstrations, number of regular and irregular government transfers	Alesina and Rodrik (1994), Persson and Tabellini (1994), Alesina and others (1996), Isham, Kaufmann, and Pritchett (1997)

Barro's political instability measures: objective	Barro and Wolf (1989), using Banks (various issues), and Barro and Lee (1994) for wars measure	Average 1960–85, or subsamples	*Political instability* Counts of revolutions; coups and revolutions per year, assassinations per million population a year; strikes *Wars* Number of years a country has experienced a war between 1960 and 1985	Barro (1991), Murphy, Schleifer, and Vishny (1991); Ojo and Oshikoya (1994); Sachs and Warner (1995), Caselli, Esquivel, and Lefort (1996); Levine and Zervos (1996), Easterly and Levine (1997)
Gupta's index of sociopolitical instability: objective	Gupta (1990), using data from Taylor and Jodice (1983)	1960–82, 104 countries	*Political instability* Number of political demonstrations, riots, political strikes, deaths from political violence, assassinations, armed attacks, political executions, successful coups d'état, unsuccessful coups d'état, nature of government	Hassan and Sarna (1996)
Alesina and others' measures of executive turnover objective	Alesina and others (1996), with underlying data from Taylor and Jodice (1983) and Banks (various issues)	1950–82, annual, 113 countries (42 African countries), some countries start in 1960	*Propensity for government change* Total government changes, major government changes (regular and coups), and irregular changes (successful coups), the propensity is estimated by regression analysis	Alesina and others (1996)
Alesina and Perotti's index of sociopolitical instability objective	Alesina and Perotti (1996), using data from Taylor and Jodice (1988)	See above	*Sociopolitical instability* Political assassinations, violent deaths per million population, successful coups, unsuccessful coups, democracy	Alesina and Perotti (1996), Perotti (1996)
Collier's civil war indicators objective	Collier (1999), based on data in Singer and Small (1994)	1816–1992, 92 countries (13 African countries)	*Duration of civil wars (monthly)* Defined on three criteria military action, active government participation, and effective resistance by each side	Collier (1999)

a In most cases, the institutional indexes are constructed as simple (unweighted) averages of all or some of the underlying components and are ordinal indexes Gastil's indexes of political and civil liberties are weighted subjectively, but these weights are not reported. For the following indexes, the unknown weights are assigned by techniques such as factor analysis Putnam's indexes, the Adelman-Morris index, Gupta's index, and the Alesina and Perotti index Bates's measures of political transition employ Gutman scales In some cases, simple counts are used Collier's civil war indicators, Barro's political instability measures and colonial measures, De Vanssay and Spindler's constitutional indicators, Taylor and Jodice's and Banks's political instability measures, and Clague and others' political regime duration measures.

based on factual observations and economic data. Risk indicators typically comprise a weighted mix of both types of measures (for example, see *Euromoney* issues of the 1990s). The table also tries to assess the coverage of Africa for these indexes, although in most cases the data on Africa are very limited.

Important Concerns about Institutional Measures

As indicated earlier, adjustments to control for problems of endogeneity and ordinality are appropriate.

Endogeneity

The important distinction between endogenous and exogenous variables is described in box 1. In general, one would expect a country's institutional structure to remain the same over time, in which case institutional variables might be considered fairly exogenous to growth. In many developing countries, however, institutional quality periodically can deteriorate sharply as a result of political instability, terms of trade or climate shocks, policy reversals, or even fiscal austerity programs. In this case, far from being exogenous, institutions may deteriorate in periods of low growth.

More alarming, subjective measures of institutions provided by the risk-rating agencies and widely used in the literature to capture economic and political efficiency may be influenced both by recent measures of growth (Haque, Mark, and Mathieson 1996) and by political events (Brewer and Rivoli 1990). Moreover, these indexes may be subject to biases through herd effects and hysteresis, meaning that bankers' judgments are too optimistic (or too pessimistic) for long periods (Somerville and Taffler 1995).

Ideally, to reduce endogeneity problems, institutional quality should be measured at the beginning of the period on which the research is concentrating. For example, if growth is averaged over 10 years, the institutional variables should be measured before—or at the beginning of—the decade. In the studies considered here, few authors take the question of endogeneity seriously. Frequently, they use indexes of institutional quality that are subjective ratings of risk compiled by private firms such as Business Environment Risk Intelligence (BERI) and International Country Risk Guide (ICRG). Typically, researchers use the endogenous middle-of-the-period ICRG index for 1982 or its components for 1960–90 or 1970–90 cross-country growth regressions (for example, Knack and Keefer 1995). The problem with the earlier BERI index for 1972 is that it covers few countries and thus reduces the samples significantly, with resulting insignificance for institutional measures (Barro 1996a).

Ordinality

Many of the institutional indexes used in table 1 are ordinal indexes. An ordinal index ranks countries on some criterion without specifying the degree of difference between countries and associates a number with the rank position (that is, 2 is second). To be used meaningfully in a growth regression, however, such an index needs to be transformed into a cardinal index, which is an index where the degree of difference matters, not just the ordering. There is no reason to presuppose that the transformation from an ordinal to a cardinal index should be one-for-one (that is, linear): for instance, the difference in the quality of the judiciary in the United States and South Africa may be much smaller than that between South Africa and Zaire, even though the same differential is measured on an ordinal scale of 1 to 10. Such possible nonlinearities, however, can be addressed using various techniques (for example, see Barro 1996a).

A separate vexing question is the often-arbitrary aggregation of different components of many of the indexes. Typically, components are simply added or averaged with the same weights (see table 1). When there are many components, factor analysis—a technique that aggregates components with unknown weights—is a convenient and superior alternative (Temple and Johnson 1998) At the least, the weighting assumptions employed should be tested (as in Knack and Keefer 1995).

Correlation with Growth

Correlations between institutional variables and growth suggest a relationship but not the direction of causality. Causality can run in both directions, from good institutions to growth and from improved growth to better institutions. Given the poor data for African countries, the relative state of African institutions can be gauged from credit-rating measures, where coverage is fairly complete and some intra-African comparisons are possible.

Simple Correlations between Institutions and Growth

The matrix in table 2 shows the degree of correlation between real per capita growth and a range of empirical institutional variables. The institutional variables are organized under headings that correspond to those in table 1. The same variables are listed in the left column and along the top. Where a column and a row meet, the statistic presented is the degree of correlation between the two variables. The shaded squares in the matrix suggest that the correlation is statistically significant at the 1 percent level. For instance, ethnolinguistic fractionalization (or ELF), measured in 1960 and entered under the heading of social characteristics, has a (statistically

Table 2. *Correlation Matrix for a Range of Institutional Variables and Real Per Capita Gross Domestic Product Growth, 1980s*

Variable	(1) *GYP80*	(2) *BERI70*	(3) *BERI80*	(4) *BUREAU*	(5) *CORRUP*	(6) *ICRG*	(7) *TELEPH*	(8) *PURGE70*
Number of countries in sample	112	42	44	63	101	101	106	100
Growth 1980–89		BERI70						
(1) Real per capita growth	1.00							
Institutional quality			BERI80					
(2) BERI index, 1970–79[a]	0.23	1.00		BUREAU				
(3) BERI index, 1980–89[a]	0.37	0.95	1.00		CORRUP			
(4) Bureaucracy index, 1980–83[a]	0.33	0.81	0.88	1.00		ICRG		
(5) Corruption index, 1980–89[a]	0.37	0.86	0.89	0.87	1.00		TELEPH	
(6) ICRG index, 1980–89[a]	0.46	0.85	0.88	0.81	0.87	1.00		PURGE70
(7) Log (telephones per 1,000 workers), 1970–79	0.33	0.83	0.84	0.82	0.71	0.75	1.00	
Social capital variables								
(8) Government purges, 1970–79	–0.02	–0.23	–0.29	–0.38	–0.26	–0.22	–0.07	1.00
(9) Antigovernment demonstrations, 1970–79	0.14	0.33	0.30	0.23	0.22	0.31	0.31	0.05
Social characteristics[b]								
(10) Ethnolinguistic fractionalization, 1960[c]	–0.30	–0.25	–0.35	–0.32	–0.36	–0.32	–0.52	–0.06
(11) Racial tension, 1984[d]	0.23	0.44	0.48	0.43	0.47	0.48	0.59	–0.03
(12) Home language not official language (percent)	–0.23	–0.49	–0.54	–0.40	–0.40	–0.37	–0.57	0.01
(13) *Minorities at risk*, 1990 (percent)	–0.16	–0.07	–0.22	–0.16	–0.17	–0.24	–0.44	–0.08
(14) Separatist movements, 1975 (percent)	–0.16	–0.21	–0.30	–0.36	–0.27	–0.25	–0.48	–0.08
Characteristic of political institutions								
(15) Constitutional change, 1970–79[e]	–0.09	–0.01	–0.20	–0.20	–0.19	–0.21	–0.18	–0.01
Political instability								
(16) Cabinet changes, 1970–79[f]	–0.03	–0.29	–0.37	0.00	–0.08	–0.13	–0.03	0.17
(17) Civil war, 1970–79[g]	–0.08	–0.39	–0.27	–0.19	–0.20	–0.32	–0.23	0.08
(18) Revolution, 1970–79[h]	–0.23	–0.28	–0.37	–0.11	–0.28	–0.38	–0.20	0.31

Note Simple correlation coefficients in the shaded squares are significant at the 1 percent level Note that the sampl sizes differ

a. The indexes are the BERI indexes (Knack 1996), Mauro's BI bureaucracy composite index (Mauro 1995), and th ICRG composite index (Knack 1996) The corruption index is from Knack and Keefer (1995) and is similar to th Mauro BI corruption index for 1980–83

b Social characteristics variables stress ethnic differences

c Ethnolinguistic fractionalization is on a scale of 0 to 1 (increasing)

	(9) *EMO70*	(10) *ELF60*	(11) *RACIAL*	(12) *HOMELA*	(13) *MINORI*	(14) *SEPAR75*	(15) *CONSTC*	(16) *CABCHG*	(17) *CIVWAR*	(18) *REVOL*
	100	107	95	111	75	70	100	99	112	100
EMO70	**1.00**									
ELF60	0 02	**1.00**								
RACIAL	0 05	**–0.71**	**1.00**							
HOMELA	–0 13	**0.76**	**–0.55**	**1 00**						
MINORI	–0.11	**0.51**	**–0.51**	**0.45**	**1 00**					
SEPAR75	–0 10	**0.62**	**–0.62**	**0.53**	**0.56**	**1 00**				
CONSTC	–0.13	0 18	–0 12	0 21	0 06	0 26	**1 00**			
CABCHG	–0 06	–0 09	0 02	0 00	–0 03	0 09	–0 03	**1.00**		
CIVWAR	0 02	0 20	**–0.34**	0 11	0 18	0 22	0 02	–0 19	**1 00**	
REVOL	–0.05	0 03	–0 12	–0 02	0 16	0 13	**0.53**	0 17	0 14	**1 00**

d Racial tension is on a scale of 0 to 6 (decreasing)

e Constitutional change captures the number of basic alterations in a constitutional structure or a new constitution

f Cabinet change refers to the number of times a premier or more than 50 percent of the cabinet is replaced

g Civil war is a dummy variable 1 if war occurs, 0 otherwise

h Revolution is an illegal attempt to replace a government

Source Author's calculations and categories using data from Easterly and Levine (1997) and Mauro (1995)

significant) correlation of minus 0.3 with real per capita growth (averaged over 1980–89).

MEASURES OF INSTITUTIONAL QUALITY. The matrix shows that some earlier variables of institutional quality (the BERI 1970–79 and the Business International—BI—bureaucracy 1980–83 indexes) are positively correlated with subsequent growth (averaged over 1980–89). The degree of correlation, of course, increases when the later (more endogenous) measures are used (the BERI 1980–89, ICRG 1980–89, and corruption 1980–89 indexes).[5]

Also of interest here are the relations *among* institutional variables, which may reveal some of their more indirect effects on growth. The number of telephones per 1,000 workers (Canning and Fay 1993) might be considered a measure of public or private service provision (public in Africa) akin to the institutional measures in Putnam (1993). Interesting positive correlations are apparent with indicators of institutional quality and social capital, and negative correlations are associated with various measures of social characteristics indicating racial diversity.

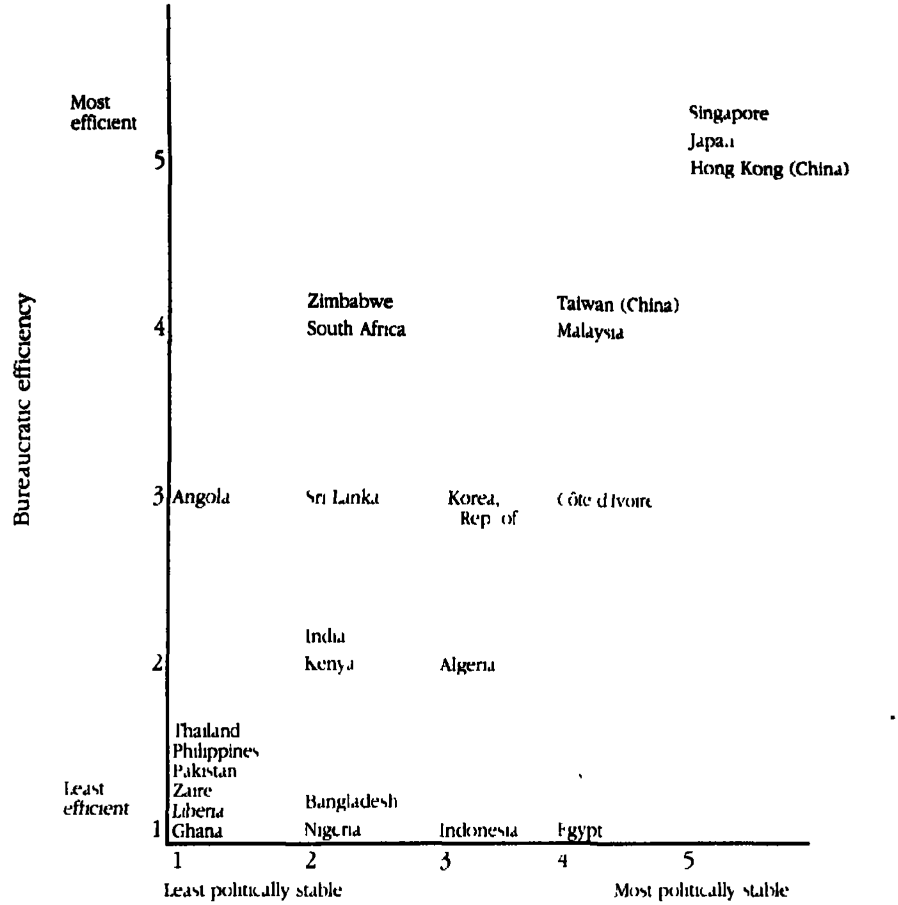
Poor institutional quality in many cases also is significantly correlated with increased political instability. The correlation between low bureaucratic efficiency and increasing political instability in African and Asian economies is shown in figure 1. It is clear that the few African countries represented in the figure lie at the bottom end of the efficiency spectrum, and some countries may have moved even farther down the scale since 1983.

MEASURES OF SOCIAL CAPITAL. Few measures of social capital are included in table 2, and the construction of such measures should prove to be a growth industry in coming years. Two variables proxy for social capital in the table: first, a measure of systematic state purges of the ranks of the regime or of the opposition, which curtail civil liberties and probably inhibit the formation of social capital, and second, peaceful antigovernment demonstrations of 100 persons or more. There are weak positive correlations between social capital and growth, while there are positive and significant correlations with some measures of formal institutional performance (the bureaucracy index).

MEASURES OF SOCIAL CHARACTERISTICS. Several measures of social characteristics reflect ethnic differences. They are highly correlated with one another, and increased ethnic diversity is negatively and significantly correlated with institutional quality and service provision. Increased ethnic diversity appears to have a negative, though largely insignificant, impact on growth. However, there is unlikely to be an endogeneity problem here because ethnicity is expected to change very slowly over time.

ATTRIBUTES OR CHARACTERISTICS OF POLITICAL INSTITUTIONS. The data are rather richer for attributes or characteristics of political institutions, and the African politi

Figure 1. *Mauro's Measures of Bureaucratic Efficiency and Political Stability African and Asian Countries, 1980–83 Averages*



Note The political stability index is the simple average of six Business International indexes institutional change, social change, opposition takeover, stability of labor, relationship with neighboring countries, and terrorism The bureaucratic efficiency index is the simple average of three Business International indexes judiciary system, red tape, and corruption

Source Adapted from Mauro (1995)

cal data set being compiled by Bates and others (1996) will prove invaluable. The example given in table 2 is constitutional change, which captures the number of alterations in a state's constitutional structure or, at the extreme, a new constitution. It is not surprising that the creation of a new constitution, which has little to do with the effectiveness or the enforcement of a constitution, however well designed, is not significantly correlated with growth or with institutional efficiency. However, constitutional change seems to be a consequence (or occasionally a cause) of revolution.

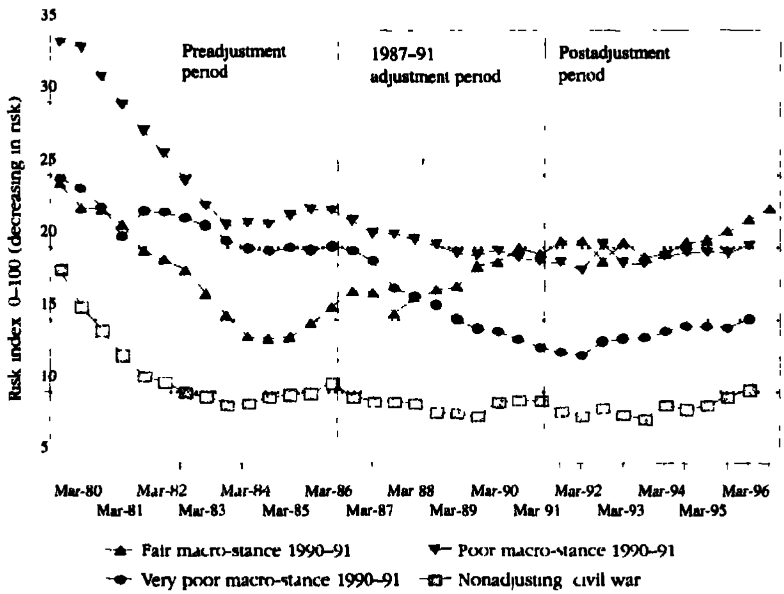
POLITICAL INSTABILITY. As for political instability, real per capita growth in 1980–89 is negatively correlated with an earlier measure of cabinet changes (1970–79) and also with two related sociopolitical measures of political instability during the 1970s—revolution and civil war.

African Economic Performance and Risk Ratings

Annual or twice-yearly credit ratings that serve as rough proxies for the enforcement of property rights are presented in figures 2–4. These data have excellent coverage of Africa. However, it should be noted that recent growth has an important bearing on these ratings.

In figures 2 and 3, African countries are categorized into four groups, each reflecting different relative economic performance. Figure 2 uses the *Institutional Investor* credit ratings, which measure the risk of default on sovereign debt (see table 1). The

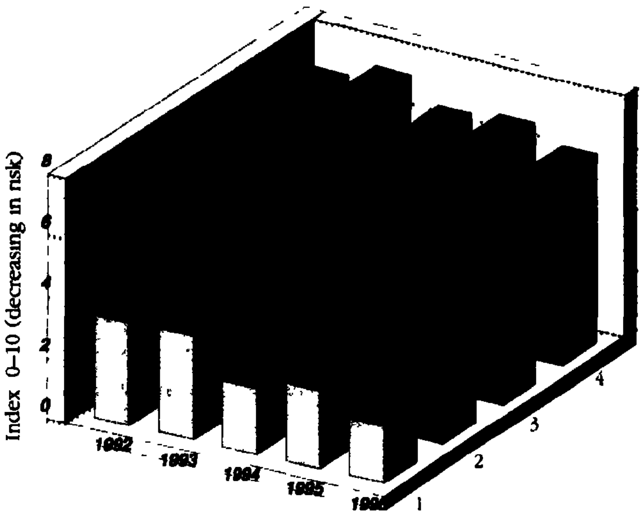
Figure 2. *Institutional Investor Risk Ratings for African Countries, 1980–96*



Note Simple averages Years in brackets after particular countries indicate the first available data point there are thus small step effects in 1993–94 Countries are categorized as follows *Fair* Burkina Faso (1993) Gabon, Ghana (1992), Guinea (1993), Kenya, Malawi (1981), Mali (1993), Nigeria, Senegal, Togo (1994) Uganda *Poor* Benin (1993), Tanzania, Zimbabwe *Very poor* Cameroon (1992), Congo, Côte d'Ivoire Mozambique (1988), Sierra Leone, Zambia *Civil war* Angola (1992), Ethiopia, Liberia, Sudan, Zaire

Source Author's calculations using data from *Institutional Investor* and country categories from World Bank (1994)

Figure 3. *Euromoney "Political Risk" Rating for African Countries, by Categories of Performance, 1992–96*



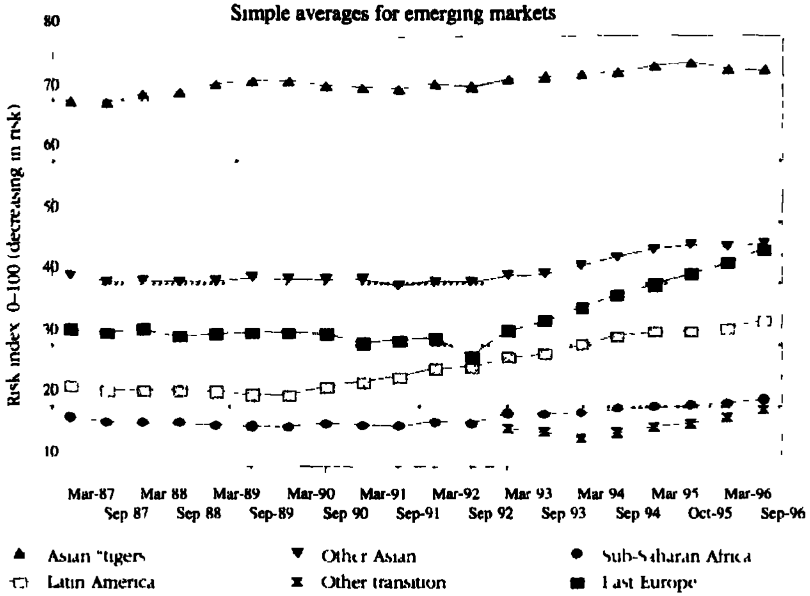
Note Weighted averages using 1993 gross domestic product Countries are categorized as follows 1 Countries with civil war in the early 1990s Angola, Burundi, Ethiopia, Liberia, Rwanda, Sierra Leone, Somalia, Sudan 2 Countries with inflation over 25 percent during 1992–94 Equatorial Guinea, Madagascar, Malawi, Mozambique, Niger, Nigeria, Tanzania Togo, Zaire, Zambia 3 Countries with poor resource allocation Cameroon, Chad, Congo, Guinea, Kenya, Lesotho Zimbabwe 4 Remaining countries Benin, Burkina Faso, Côte d'Ivoire, The Gambia, Ghana, Guinea-Bissau, Mali, Mauritania Senegal Uganda

Source Author's calculations using disaggregated data from *Euromoney* and country categories from Collier and Gunning (1999)

assessment covers economic performance in the 1980s from the standpoint of the overall macroeconomic situation in 1990–91, *after* a period of structural adjustment in 1987–91 (categories from World Bank 1994). With the caveat that the comparability of these survey indicators across time may be in doubt, in figure 2 the countries categorized as "fair" (1990–91) start from a low ratings base, before adjustment, and continue improving in the eyes of foreign investors during the 1990s. "Poor" and "very poor" countries steadily decline in the preadjustment and adjustment periods: the decline in both groups, however, shows signs of leveling off in the 1990s. Countries involved in civil war show little improvement, although Ethiopia's rating increased marginally in the mid-1990s.

Figure 3 examines political risk, the one subjective component in the *Euromoney* rating of aggregate risk. The assessment reviews economic performance in the 1990s from the standpoint of the country's macroeconomic position in 1992–94 (categories from Collier and Gunning 1999). Countries are ranked by four "hurdles" and show distinct ratings by economic performance (but the resource allocation criterion proves a weak means of distinguishing countries).

Figure 4. *Institutional Investor Risk Ratings for Africa and Emerging Markets, 1987–96*



Note Countries are categorized as follows *East Europe* Czech Republic, Hungary, Poland, Romania, Slovakia Slovenia *Other transition* Albania, Croatia Estonia, Georgia Kazakhstan, Latvia, Lithuania, Russia, Ukraine Uzbekistan, Yugoslavia *Asian tigers* Hong Kong (China), Republic of Korea, Singapore, Taiwan (China) *Other Asian* Bangladesh, China, India, Indonesia, Malaysia, Pakistan, Philippines, Sri Lanka, Thailand *Latin America* Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica Ecuador El Salvador, Guatemala, Honduras Mexico, Paraguay, Peru, Uruguay Venezuela *Africa* Angola, Benin, Burkina Faso, Burundi, Cameroon, Chad Congo, Côte d Ivoire Equatorial Guinea, Ethiopia, The Gambia Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique Niger, Nigeria, Rwanda, Senegal, Sierra Leone Somalia, Sudan, Tanzania, Togo, Uganda, Zaire, Zambia, Zimbabwe

Source Author's calculations using data from *Institutional Investor*

From Correlation to Causation

The evidence for causation in the literature on growth and institutions may be compromised not only by deficiencies in the institutional measures but also by the poor quality of the data for other variables. Many growth regressions use data from the cross-country Penn World Tables data set (Summers and Heston 1991); almost half of the 138 countries in this set have poor-quality data, and of these, 37 are African countries.

How Do Institutions Affect Growth?

The usefulness of North's framework, in conjunction with neoclassical growth models, is that it helps to clarify the likely channels of influence on growth of measures o

institutional quality and social capital (proxies for the performance of formal and informal institutions). Box 2, on page 102, suggests how North's institutional framework could be integrated with Solow's growth model (see also Tornell 1997).

Reference has already been made to North's contention that weak institutions may have a direct effect on growth because they reduce the efficiency of investment. The efficiency effect might arise from the initial state of institutions in the model, which will affect the productivity of the existing stock of capital, or it may result from changes in institutions over time. In an environment where the enforcement of property rights is not reliable, firms will tend to be small scale, to use low-capital technology, and to have short-term horizons. Second, there could be an indirect effect on growth resulting from a decline in investment—a variable that is already in the Solow model—such as increased transaction costs through bribes and rent seeking.

Most studies in the literature have used reduced-form growth models in which the direct and indirect effects of institutions are difficult to disentangle. Recently a few authors (such as Mauro 1995) have used Solow-type models as well as separate investment models to evaluate the role of the quality of institutions, testing the extent of social capital and the effectiveness of the laws on property rights both for their direct effects on growth (more efficient investment) and for their indirect effects larger volume of investment).

The three remaining categories of institutional measures in table 1—political characteristics, political instability (duration), and social characteristics—are often included n growth and investment models as proxies for institutions. However, the observation hat measures of political instability are only weakly significant—or insignificant—in growth regressions when the quality of institutions is controlled for, while evidence for "democracy" measures is generally weak in growth regressions, has led authors such as Clague and others (1996) to use these measures instead to model property rights. Similarly, Barro (1996b) recently suggested using indicators of political stability as a determinant of property rights. Finally, Barro (1996a) suggests that what may influence the evolution of a political regime itself are persistent social characteristics such as ethnic diversity, religious or colonial heritage, and past growth

Findings from Models of Growth

Most of the institutions and growth studies considered here are cross-sectional regressions, that is, they use data from many countries for one period only (which may be an average of years). Increasingly, however, growth analysts are turning to panel data, which cover a range of countries for two or more consecutive periods (ideally, the periods are averages of 7 to 10 years, close to business-cycle frequencies).[6]

ACCOUNTING FOR POLITICAL CHARACTERISTICS, POLITICAL INSTABILITY, AND SOCIAL CHARACTERISTICS. Studies that use objective measures of democracy (that

is, attributes, not performance) tend to find inconclusive results in growth regressions because democracy, acting through various channels, may have both positive and negative implications for growth (Alesina and Rodrik 1994; Helliwell 1994; Alesina and others 1996). On the positive side, transparency and accountability may enhance economic and other rights, including respect for contracts. On the negative side, the consensus required by democratic institutions, or interest group lobbies, may delay responses to shocks and implementation of legislation. A problem with using a single cross-section is that it is difficult to distinguish between changes, their different durations, and the different kinds of dictatorships. Growth performance may be poor during periods of democratic transition because the government must confront the economic legacy of a collapsed regime and the attendant political instability (Alesina and Perotti 1994). Finally, democracy tends to be correlated with high-income countries, good institutions, and an educated workforce. The effects are difficult to disentangle.

Variables used to indicate political instability generally suffer from endogeneity problems. Although, for example, political instability may reduce the volume and efficiency of investment and thus hamper growth, negative income shocks may also promote political instability. Alesina and others (1996) model the effects of executive instability by first measuring the probability of a change in government for three definitions of change (every change in government, major changes in government, and coups d'état) and then estimating the relationship between growth in per capita income and a range of variables, including the probability of a change of government. They find apparently robust results indicating that high executive turnover has a negative impact on growth (but see a methodological critique in Deaton and Miller 1996). Bienen, Londregan, and van de Walle (1993), however, examine the reverse causality and find that current and lagged growth rates inhibit the transfer of power. Londregan and Poole (1990) assume, plausibly, that the current rate of growth and the probability of a coup are not simultaneously linked but also, perhaps less plausibly, that past coups have no effect on current income. They too find that income growth inhibits coups.

Simpler counts of revolutions and coups appear to influence growth negatively, but these effects disappear when property rights are controlled for (Barro 1996a) Aggregate indexes of riots, demonstrations, and assassinations do not appear to affect growth in these estimates, but they may reduce investment and saving (Alesina and Perotti 1996). Here they may proxy for ineffective property rights Svensson (1998) finds a link between political instability and the quality of property rights. He also finds that when property rights are included in the investment equation, the political instability variables cease to be significant.

Similarly, Clague and others (1996) find that the characteristics and stability of political regimes (type and duration) appear to be important determinants of the quality of economic institutions. This research highlights the tradeoff between credibility and executive flexibility as a possible explanation for why some democratic

regimes do not deliver the goods, while some longer-lived autocratic regimes do. Using annual data to account for frequent changes of regime, the paper finds that short-lived democracies are least likely to ensure adequate property rights, while longer-term democracies offer better protection for property and contract rights than any other type of regime of any duration.

Barro (1996a), who evaluates with panel data the role of social characteristics in determining political regimes using proportions of the population by religion, rates of urbanization, colonial heritage, ethnic differences, income inequality, and social indicators such as life expectancy, actually forecasts which countries are likely to experience declines and improvements in future democracy. That none of the above variables is significant (except an indicator for countries in the Organization of Petroleum-Exporting Countries) may be related to Barro's neglect of the *duration* of democratic and other regimes, which Clague and others (1996) emphasize. Several of these variables also have been shown to be relevant in Solow-type growth equations, although with doubtful robustness (Temple 1999).

ACCOUNTING FOR SOCIAL CAPITAL AND QUALITY OF INSTITUTIONS. In early work examining the influence of institutions on growth, Kormendi and Meguire (1985) and Scully (1988) used cross-sectional growth regressions and Gastil's indexes of civil and political rights. In these two studies, Gastil's indexes are transformed so as to address the ordinality problem. Kormendi and Meguire's paper suggests an indirect effect on growth through investment, but because the measure of civil rights is measured contemporaneously with growth and other variables, endogeneity problems arise, and no definite causal relation is established. Scully's study examines whether civil, political, and economic rights have a direct efficiency-enhancing effect on growth, and again, the evidence is mixed and is compromised by endogeneity problems.

More recently, Isham, Kaufmann, and Pritchett (1997) find that purely political rights (a measure of the quality of democracy) and type of political regime (attributes) have no effect on the rates of return to World Bank–financed government investment projects but that civil rights are significant—which we classify as a possible "social capital" proxy in table 1. However, the indexes are both ordinal and endogenous.

An interesting paper by Knack and Keefer (1997) examines informal institutions using measures of trust and civic norms drawn from the World Values Surveys (see Inglehart 1994a). The authors find that in reduced-form growth regressions, singly and together, trust and civic norms are positively associated with growth and claim a causal role. Specification tests for the robustness of these results find that they are fairly insensitive to changes in specification, exclusions of influential observations, and the use of additional regressors. In the Solow regressions (not reported), these variables lose significance or become insignificant. This could suggest that the in-

vestment relationship where both variables, but especially civil norms, prove to be important fully captures their (indirect) impact on growth. However, as with other studies surveyed here, the endogeneity of these ordinal indexes and their measurement error raise doubts about a robust causal role for these measures. There probably is considerable measurement error here: some groups are oversampled (higher-status groups), which the authors try to correct for, but the five survey questions comprising the civic performance measure may not have been answered truthfully, while the question regarding trust is ambiguous.

The authors also test another proxy for social capital (following Putnam 1993): associational activity or membership in groups, again using World Values Surveys data. They find no significance for these measures and perverse relations in some cases, but this measure is flawed because it fails to reflect the intensity (or quality) of membership. Thus the authors' conclusion that declining social capital reflects declining trust and civic norms rather than declining associational life may be too strong.

Finally, Knack and Keefer (1997) examine the impact of trust and civic norms on the quality of formal institutions and report positive and significant links, although it should be noted that there are endogeneity problems. They also find evidence that income equality and low ethnic divisions are linked with trust and civic cooperation.

Clague and others (1995) use an innovative objective measure of contract enforcement that they call contract-intensive money for 96 countries during 1969–90. They describe this stock of money as the amount that is held on deposit in banks and other financial institutions. They reason that where third-party enforcement of contracts is not reliable and insecure property rights hamper or preclude the use of assets as collateral for loans, banks and financial intermediaries will not profit from providing low-cost retail banking services or from offering incentives to attract deposits. In such cases, where there are few advantages to holding money on deposit and customers also face the risk of nonrecovery, individuals will hold a lower proportion of their assets in such accounts. The authors are at pains to show that this measure is not just an indicator of financial development but that it is also highly significant in an investment equation (controlling for financial depth), suggesting an indirect influence on growth.[7] They fail to find evidence for a direct or efficiency influence on growth.

The remaining studies all use data from the risk-rating agencies referred to earlier to assess the effects of institutional variables on investment and growth: the ICRG, BERI, and BI. They generally show that measures of corruption and weak institutions have the expected negative effect on investment and growth, while efficient bureaucracy and observation of the rule of law effectively support the achievement of these objectives.

Mauro (1995) reports results from three regressions: an investment equation, a reduced-form growth regression, and an extended Solow growth equation. As expected, he finds that efficient bureaucracy (although not corruption) indexes are significant in an investment equation (and hence influence growth indirectly) and

demonstrates their importance in a broad range of simple correlations with investment for different periods and different types of investment. In reduced-form regressions the bureaucracy index is robustly significant, but corruption is insignificant. In a Solow model bureaucracy is barely significant, and corruption is insignificant. With a broader set of variables, including political instability measures, the significance of bureaucracy diminishes further. Mauro attempts to address the endogeneity problems with both the institutional variables and investment, but in neither case is the institutional index significant. It is surprising that the Solow model lacks a robust institutional "efficiency" link with investment, but this absence may be due to the ordinality, aggregation, and endogeneity problems discussed earlier.

Helliwell (1996b) is one of the few studies where the institutional variables are arguably not too endogenous, since a bureaucracy index for 1981–83 for some Asian countries is related to average growth for the subsequent decade. In a reduced-form growth equation, the institutional linkages surprisingly do not show up (investment is absent, and the institutional variables are "free to take all the credit"). The answer may lie with measurement and ordinality problems.

Knack and Keefer (1995) and Knack (1996) use two institutional indexes in growth regressions capturing security of contract and property rights, the ICRG index from 1982, and the earlier and hence less endogenous BERI index (from 1972). They report some results without statistics to support them. It appears that both institutional indexes are significant in investment regressions during 1974–89 and 1960–89, confirming an indirect effect on growth through factor accumulation. There also appears to be weak evidence for a direct efficiency effect on growth.

Lane and Tornell (1996) address two commonly omitted variables—natural resource abundance and institutions—and attempt to explain why some resource-rich countries have lower growth rates than resource-poor countries. They report reduced-form growth and investment regressions, and their results focus on the effects of weak property rights (reflected in estimates of the ICRG risk indicator) and an increase in manufacturing concentration. The coexistence of weak institutions and powerful industrial groups affects growth and investment negatively and significantly.

A 1960–90 panel study by Barro (1996a) uses a Solow model and the ordinal ICRG and the earlier BERI indexes of property rights. With an endogenous rule of law index (dated 1982), he finds a consistently positive and significant effect on growth. But, unfortunately, probably due to reduced sample size using the less endogenous BERI index, it is rendered only marginally significant (although not too different in value). In the same study, Barro transforms Gastil's political rights index into three categories—high, middle, and low measures of the quality of democracy—and finds that the middle level of democracy most favors growth. The implication is that where a moderate amount of political rights already has been extended, a further increase could actually diminish growth, perhaps due to pressures for income redistribution.[8]

What We Know and What We Don't

The more recent literature suggests that the appropriate institutional variables to include in investment and growth regressions are those that capture the *performance* or *quality* of formal and informal institutions rather than merely describe the characteristics or attributes of political institutions and society or measure their political instability. Reinforcing this view, evidence suggests that the characteristics and duration of political regimes are important in determining whether countries have well-enforced and well-defined property rights, demonstrating a possible link between political stability and investment and implying that inefficient political institutions with high transaction costs may result in weak property rights and thus discourage economic growth. Barro (1996a) has attempted to model such political characteristics as a function of cultural, social, and historical variables (including indicators for colonialism), although not with particularly conclusive results.

The performance or quality measures for formal and informal institutions include respect for contracts, property rights, trust, and civil freedom. Evidence suggests that the quality of institutions has a robust and significant indirect relationship to growth via its effect on the volume of investment. There also is evidence, although it is weak, for a direct relationship between institutions and growth.

Thus better-performing institutions may improve growth by increasing the volume of investment—for example, by eliminating bureaucratic red tape and rent-seeking costs—and (more weakly) by improving the efficiency of investment, say, by enforcing well-defined property rights. Similarly, the promotion of social capital strengthening informal institutions may positively influence growth both directly and indirectly.

It is important to highlight the simultaneous relationship between growth, investment, and institutions. However, most of the studies in this survey ignore simultaneity issues and have other methodological problems found in cross-country growth studies in general (see Temple 1999). In particular, they often deal inadequately with endogenous institutional measures, while there are aggregation biases and other problems with the institutional measures used. Thus a definite positive conclusion on the links between growth and institutions is difficult to pin down, suggesting that the claims for causality should be treated with caution.

Although the results are suggestive, many commentators are skeptical because of the problems that plague cross-country studies. However, it is possible to take a more constructive view of this literature, from which a great deal already has been learned. Recent literature is paying increasing attention to many of the methodological difficulties, and the result is more thorough specification and other tests, more attention to endogeneity issues, and more guarded and less cavalier policy conclusions. There are limited data in the institutional sphere, but the construction of

more subtle empirical measures for institutions is by now a growth area in itself. Admittedly, the growth and institutions models are highly stylized, but this frontier research area may well yield stronger conclusions in the future.

To the extent that these tentative results are reinforced by later, more careful research, this would underline the importance of the state in facilitating the development and enforcement of an independent and effective judiciary, in refraining from predatory actions that discourage saving, investment, and production, and in extending civil and political rights that promote the development of social capital.

Although the limited evidence reveals the likelihood of linear relationships between various efficiency measures of institutions and growth, Barro (1996a) finds evidence suggesting that after a certain threshold, further extension of political rights could retard growth, perhaps due to pressures for income redistribution, and even argues against the desirability of exporting democratic institutions to developing nations (for example, countries such as Mexico and Malaysia). This conclusion probably is overstated, and another interpretation can be considered. It is possible that Barro is picking up strong effects from a subset of authoritarian and relatively recently fast-growing countries that have exploited the "catch-up" phase and increasing returns to scale and now are entering a mature period of slower growth. Increased demand for democracy could be due to the consequences of slowing growth (for example, recent trade union demands in the Republic of Korea) as well as to exposure to democratic rights through internationalization and higher standards of living. However, it also is possible that countries like South Africa that have a high potential for catch-up and have recently extended political rights may suffer delays in implementation and setbacks to economic reform under participatory politics.

The complex area of institutional change lies outside the scope of this review. A crucial problem for developing countries is to achieve institutional credibility via third-party enforcement. Even if a neutral state is able to monitor property rights and enforce contracts effectively, the question arises as to who guards the guards. Many observers are skeptical that constitutional reform will be capable of restraining the tyrannical exercise of power. Others believe that the permanent extension of civil and political liberties may serve to restrain the state.

Clearly it is important initially to match the state's role to its capabilities as well as to foster growth by invigorating institutions In the context of law reform, Posner (1998) believes that fairly modest fiscal expenditures may secure reforms that could enhance economic growth, thereby generating further resources to proceed with deeper reforms, and so on, in what he describes as a virtuous circle. Other authors agree that a less costly and more rapid reform involves enacting efficient rules to be administered by less-than-efficient institutions, rather than wholesale, expensive, and time-

consuming reforms of the institutions themselves (Hay, Schleifer, and Vishny 1996). However, a more sophisticated legal reform will be required to protect civil and political liberties.

Whether such profound institutional change and development can be rapidly achieved and sustained is uncertain. Institutionalists tend to emphasize the slow pace of change Indeed, North's sobering view is that "creating a system of effective enforcement and of moral constraints on behavior is a long, slow process that requires time to develop if it is to evolve—a condition markedly absent in the rapid transformation of Africa from tribal societies to market economies" (North 1990: 60).

Notes

Janine Aron holds a research fellowship from the Economic and Social Research Council, U K., at the Centre for the Study of African Economies, Institute of Economics and Statistics, Oxford University This article was prepared as a background paper for the World Bank's *World Development Report 1997· The State in a Changing World* The author would like to thank Dennis Anderson, Robert Bates, David Bevan, Stephen Ellis, John Knight, John Muellbauer, Avner Offer, Luis Servén, Jonathan Temple, Nicolas van de Walle, and members of the Centre for the Study of African Economies for comments.

1. Recent evidence on Africa's growth performance is recorded in Collier and Gunning (1999), Easterly and Levine (1997), Elbadawi and Ndulu (1995), and Sachs and Warner (1997), while Ravallion and Chen (1997) provide a recent analysis of poverty and distribution

2 Characteristics of political regimes often are called measures of democracy By contrast, Bates and others (1996) stress that their scales of detailed features of political institutions do not measure "democracy". they measure the attributes rather than the performance of institutions. (Such measures do clarify the scope for political reform, however)

3 Alesina and Perotti (1994) categorize variables as definitions of democracy (civil and economic liberties and competitiveness of elections) and political instability (sociopolitical instability, executive turnover, and the risk-agency measures of corruption and bureaucracy) Brunetti (1997) categorizes political variables into democracy, government stability, political violence, policy volatility, and subjective perceptions of politics

4 Gastil's (1991) subjective rankings of civil liberty and political liberty have been constructed annually since 1973 for a large number of countries using various public sources, including newspaper reports, human rights reports, and U.S State Department reports to Congress on human rights in countries that receive U.S aid (see Scully 1987 for further discussion) Political rights rankings are based on criteria reflecting the degree to which citizens have control over government, while civil rights rankings use criteria reflecting individual rights relative to the state (see the criteria in table 1)

5 An examination of the underlying data shows that in practice the ICRG measures of Knack and Keefer (1995) are predominantly measured *later* in the 1980s, that is, 1984/85–89 (this is especially true for African countries)

6 It remains controversial whether panel studies can better address the vexing problem of endogenous regressors For discussion on this, see Caselli, Esquivel, and Lefort (1996) and Arellano and Bover (1995) Moreover, Pritchett (1998) argues against the usefulness of panel data for investigating long-term growth rates in developing countries, given the great instability in their growth rates over time

7 The authors find that the results are somewhat sensitive to the exclusion of outliers but find no evidence of reverse causality in the measure of contract-intensive money

8. Given the very limited variation in the institutional indexes for African countries, often with serious endogeneity problems, there probably is little to gain on the role of institutions from studies focused solely on Africa, such as the panel data studies of Ghura and Hadjimichael (1996) and Savvides (1995) and the cross-sectional growth study of Sachs and Warner (1997).

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Adelman, Irma, and Cynthia Taft Morris 1968 "Performance Criteria for Evaluating Economic Development Potential: An Operational Approach" *Quarterly Journal of Economics* 82 (2). 260–80

Alesina, Alberto, and Roberto Perotti 1994 "The Political Economy of Growth: A Critical Survey of the Recent Literature" *The World Bank Economic Review* 8(3):351–71

———. 1996 "Income Distribution, Political Instability, and Investment." *European Economic Review* 40(6, June).1203–28

Alesina, Alberto, and Dani Rodrik 1994 "Distributive Politics and Economic Growth" *Quarterly Journal of Economics* 109(May).465–90

Alesina, Alberto, Sule Ozler, Nouriel Roubini, and Phillip Swagel 1996 "Political Instability and Economic Growth." *Journal of Economic Growth* 1(2) 189–211

Arellano, Manuel, and Olympia Bover 1995 "Another Look at the Instrumental Variable Estimation of Error-Components Models" *Journal of Econometrics* 68(1, July) 29–51

Aron, Janine 1996 "The Institutional Foundations of Growth" In Stephen Ellis, ed, *Africa Now People, Policies, and Institutions*, ch 6 London James Currey/Heinemann, Paris Karthala

———. 1998. "Political, Economic, and Social Institutions A Review of Growth Evidence" Working Paper Series 98-4 Centre for the Study of African Economies, Oxford University Processed

Banks, Arthur, ed Various issues, 1975–1982/83 *Political Handbook of the World* Boston, Mass McGraw-Hill

———, ed Various issues, 1983/84, annually *Political Handbook of the World* New York: CSA Publications

Barro, Robert J 1991 "Economic Growth in a Cross-Section of Countries" *Quarterly Journal of Economics* 106(2, May) 407–43

——— 1996a "Democracy and Growth" *Journal of Economic Growth* 1(1) 1–27

——— 1996b "Institutions and Growth An Introductory Essay" *Journal of Economic Growth* 1(1) 145–48

Barro, Robert J, and Jong-Wha Lee. 1994 "Losers and Winners in Economic Growth" In Michael Bruno and Boris Pleskovic, eds, *Proceedings of the World Bank Annual Conference on Development Economics 1993*, pp. 267–97 Washington, D C World Bank

Barro, Robert, and Xavier Sala-i-Martin 1995 *Economic Growth* Boston, Mass · McGraw-Hill.

Barro, Robert, and Holger Wolf 1989 "Data Appendix for Economic Growth in a Cross-Section of Countries." National Bureau of Economic Research, Cambridge, Mass Processed

Bates, Robert H. 1989 *Beyond the Miracle of the Market The Political Economy of Agrarian Development in Rural Kenya* Cambridge, U K Cambridge University Press

Bates, Robert H, Karen Ferree, A J Robinson, Sumita Singh, and Anne Wren 1996 "Toward the Systematic Study of Transitions" Development Discussion Paper 526 Harvard Institute of International Development, Harvard University, Cambridge, Mass. Processed

Bienen, H. S., John Londregan, and Nicolas van de Walle. 1993 "Ethnicity, Leadership Succession, and Economic Development in Africa." Institute for Policy Reform, Washington, D.C. Processed.

Bollen, Kenneth. 1990. "Political Democracy: Conceptual and Measurement Traps." *Studies in Comparative International Development* 25(1, spring):7–24

Borner, Silvio, Aymo Brunetti, and Beatrice Weder 1995 *Political Credibility and Economic Development* New York: St Martin's Press

Brewer, Thomas, and Pietra Rivoli. 1990. "Politics and Perceived Country Creditworthiness in International Banking." *Journal of Money, Credit, and Banking* 22(3).357–69

Brunetti, Aymo. 1997 "Political Variables in Cross-Country Analysis." *Journal of Economic Surveys* 11(2, June):163–90.

Brunetti, Aymo, Gregory Kisunko, and Beatrice Weder 1998. "Credibility of Rules and Economic Growth Evidence from a Worldwide Survey of the Private Sector " *The World Bank Economic Review* 12(3)·353–84.

Canning, David, and Marianne Fay 1993. "The Effect of Transportation Networks on Economic Growth " Discussion Paper Series Department of Economics, Columbia University, New York Processed.

Caselli, Francesco, Gerard Esquivel, and Fernando Lefort. 1996 "Reopening the Convergence Debate A New Look at Cross-Country Empirics " *Journal of Economic Growth* 1(3, September):363–69

Clague, Christopher, Philip Keefer, Stephen Knack, and Mancur Olsen. 1995. "Contract-Intensive Money. Contract Enforcement, Property Rights, and Economic Performance." Economics Department, University of Maryland, College Park Processed

——— 1996 "Property and Contract Rights in Autocracies and Democracies." *Journal of Economic Growth* 1(2, June) 243–76

Collier, Paul 1996 "The Role of the African State in Building Agencies of Restraint " In J Mats Lundahl and Benno Ndulu, eds , *New Directions in Development Economics Growth, Environmental Concerns, and Government in the 1990s*, pp 282–98 Studies in Development Economics 3 New York Routledge

——— 1999 "On the Economic Consequences of Civil War " *Oxford Economic Papers* 51(1) 168–83

Collier, Paul, and Jan Gunning 1999 "Explaining African Economic Performance " *Journal of Economic Literature* 37(1)·64–111

Cukierman, Alex, Steven B Webb, and Bilin Neyapti 1992 "Measuring the Independence of Central Banks and Its Effect on Policy Outcomes " *The World Bank Economic Review* 6(3, September) 353–98

David, Paul 1994 "Why Are Institutions the 'Carriers of History'? Path Dependence and the Evolution of Conventions, Organizations, and Institutions " *Structural Change and Economic Dynamics* 5(2)·205–20

———. 1997 "Rethinking Technology Transfers· Incentives, Institutions, and Knowledge-Based Industrial Development " In Charles Feinstein and Christopher Howe, eds., *Chinese Technology Transfer in the 1990s Current Experience, Historical Problems, and International Perspectives* Cheltenham, U K : Edward Elgar

Deaton, Angus S , and Ronald I Miller 1996 "International Commodity Prices, Macroeconomic Performance, and Politics in Sub-Saharan Africa " *Journal of African Economies* 5(3, supplement, pt 1) 99–191

De Vanssay, Xavier, and Zane A Spindler. 1992. "Freedom and Growth· Do Constitutions Matter?" Economics Department, Simon Fraser University, Vancouver, B.C , Canada. Processed

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4, November):1203–50.

Elbadawi, Ibrahim, and Benno Ndulu. 1995. "Growth and Development in Sub-Saharan Africa: Evidence on Key Factors." Paper prepared for the 1995 World Congress of the International Economic Association, Tunis, December. Processed

Gastil, Raymond D. 1989. *Freedom in the World Political Rights and Civil Liberties, 1988–89* Washington, D C.: Freedom House

———. 1991. "The Comparative Survey of Freedom. Experiences and Suggestion." In Alex Inkeles, ed., *On Measuring Democracy.* New Brunswick, N J . Transaction Publishers

Ghura, Dhaneshwar, and Michael T Hadjimichael 1996 "Growth in Sub-Saharan Africa." *International Monetary Fund Staff Papers* 43(3, September).605–34.

Gupta, Dipak K. 1990 *The Economics of Political Violence The Effect of Political Instability on Economic Growth.* New York and London Greenwood Press, Praeger

Gurr, Ted. 1990. "Polity II. Political Structures and Regime Change 1800–1986." Inter-University Consortium for Political and Social Research, Ann Arbor, Mich Processed

Haque, Nadeem U., Nelson Mark, and Donald T Mathieson 1996. "The Economic Content of Indicators of Developing-Country Creditworthiness " *International Monetary Fund Staff Papers* 43(4, December)·688–724

Hassan, Fareed, and Naveen Sarna 1996 "Economic Growth and Institutions in Africa. A Cross-Country Time Series Analysis." Europe and Central Asia Department I, World Bank, Washington, D C. Processed

Hay, Jonathan, Andrei Schleifer, and Robert W Vishny 1996 "Privatization in Transition Economies: Towards a Theory of Legal Reform." *European Economic Review* 40(315):559–68

Helliwell, John F. 1994 "Empirical Linkages between Democracy and Economic Growth " *British Journal of Political Science* 24(2, April).225–48

———. 1996a. "Do Borders Matter for Social Capital? Economic Growth and Civic Culture in U.S. States and Canadian Provinces." NBER Working Paper 5863 National Bureau of Economic Research, Cambridge, Mass Processed.

——— 1996b. "Economic Growth and Social Capital in Asia " NBER Working Paper 5470. National Bureau of Economic Research, Cambridge, Mass. Processed

Helliwell, John F., and Robert Putnam 1995. "Social Capital and Economic Growth in Italy " *Eastern Economic Journal* 21(3)·295–307

Heston, Alan, and Robert Summers. 1996 "International Price and Quantity Comparisons: Potentials and Pitfalls." *American Economic Review* 86(2)·20–24

Inglehart, Ronald 1994a. *Codebook for World Values Surveys* Ann Arbor, Mich. Institute for Social Research

———. 1994b. "The Impact of Culture on Economic Development Theory, Hypotheses, and Some Empirical Tests " Department of Political Science, University of Michigan, Ann Arbor. Processed

Isham, Jonathan, Daniel Kaufmann, and Lant Pritchett 1997. "Civil Liberties, Democracy, and the Performance of Government Projects " *The World Bank Economic Review* 11(2):219–42

Knack, Stephen. 1996. "Institutions and the Convergence Hypothesis: The Cross-National Evidence." *Public Choice* 87(3–4):207–28

Knack, Stephen, and Philip Keefer. 1995. "Institutions and Economic Performance. Cross-Country Tests Using Alternative Institutional Measures " *Economics and Politics* 7(3):207–27.

———. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation " *Quarterly Journal of Economics* 112(4, November) 1251–88

Kormendi, Roger C., and Philip G. Meguire 1985. "Macroeconomic Determinants of Growth. Cross-Country Evidence." *Journal of Monetary Economics* 16(2, September):141–63.

Lal, Deepak, and H. Myint 1996. *The Political Economy of Poverty, Equity, and Growth* Oxford, U.K.. Clarendon Press.

Lane, Philip R., and Aaron Tornell. 1996. "Power, Growth, and the Voracity Effect." *Journal of Economic Growth* 1(2, June) 213–41.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Regressions " *American Economic Review* 82(4, September) 942–63

Levine, Ross, and Sara Zervos 1996. "Stock Market Development and Long-Run Growth " *The World Bank Economic Review* 10(2).323–42

Londregan, John, and Keith Poole 1990. "Poverty, the Coup Trap, and the Seizure of Executive Power." *World Politics* 42(2, January).151–83.

Mankiw, N Gregory, David Romer, and David N Weil 1992 "A Contribution to the Empirics of Economic Growth " *Quarterly Journal of Economics* 107(2) 407–37

Mauro, Paolo. 1995. "Corruption and Growth " *Quarterly Journal of Economics* 110(3, August)·681–712

Murphy, Kevin M , Andrei Schleifer, and Robert W Vishny 1991 "The Allocation of Talent Implications for Growth." *Quarterly Journal of Economics* 106(2, May).503–30

Ng, Francis, and Alexander Yeats 1999. "Good Governance and Trade Policy· Are They the Keys to Africa's Global Integration and Growth?" Trade Research Team, World Bank, Washington, D C Processed

North, Douglass 1990 *Institutions, Institutional Change, and Economic Performance* New York Cambridge University Press

Ojo, Oladeji, and Temitope Oshikoya 1994 "Determinants of Long-Term Growth Some African Results " *Journal of African Economies* 4(2) 163–91

Perotti, Roberto 1996. "Growth, Income Distribution, and Democracy· What the Data Say." *Journal of Economic Growth* 1(June) 149–87

Persson, Torsten, and Guido Tabellini 1994 "Is Inequality Harmful for Growth? Theory and Evidence " *American Economic Review* 84(3, June) 600–21

Posen, Adam 1998. "Do Better Institutions Make Better Policy?" *International Finance* 1(1)·173–205

Posner, Richard A. 1998 "Creating a Legal Framework for Economic Development " *The World Bank Research Observer* 13(1)·1–11

Pritchett, Lant 1998 "Patterns of Economic Growth. Hills, Plateaus, Mountains, and Plains " Policy Research Working Paper 1947 Policy Research Department, World Bank, Washington D C Processed

Przeworski, Adam, and Fernando Limongi 1993 "Political Regimes and Economic Growth." *Journal of Economic Perspectives* 7(3).51–69

Putnam, Robert D 1993 *Making Democracy Work· Civic Traditions in Modern Italy* Princeton, N J : Princeton University Press

Ravallion, Martin, and Shaohua Chen 1997 "What Can New Survey Data Tell Us about Recent Changes in Distribution and Poverty?" *The World Bank Economic Review* 11(2, May)·357–82

Sachs, Jeffrey D , and Andrew M Warner 1995 "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 0(1) 1–95.

———. 1997. "Sources of Slow Growth in African Economies " *Journal of African Economies* 6(3):335–76.

Savvides, Andreas 1995 "Economic Growth in Africa." *World Development* 23(3):449–58.

Scully, Gerald W. 1987. "The Choice of Law and the Extent of Liberty " *Journal of Institutional and Theoretical Economics* 143(4, December):595–615

———. 1988. "The Institutional Framework and Economic Development " *Journal of Political Economy* 96(3):652–62.

Singer, M., and J. D. Small 1994. "Correlates of War Project International and Civil War Data, 1816–1992." Inter-University Consortium for Political and Social Research, Ann Arbor, Mich Processed.

Solow, Robert M. 1956 "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70(1):65–94

Somerville, R. A., and R. J Taffler 1995. "Banker Judgment versus Formal Forecasting Models The Case of Country Risk Assessment." *Journal of Banking and Finance* 19(2, May).281–97

Spindler, Zane 1991 "Liberty and Development· A Further Empirical Perspective " *Public Choice* 69(2, February):197–210.

Srinivasan, T. N 1995. "Long-Run Growth Theory and Empirics: Anything New?" In Takatoshi Ito and Anne O. Krueger, eds., *National Bureau of Economic Research–East Asia Seminar on Economics Series*, vol. 4, pp 37–66. Chicago, Ill : University of Chicago Press

Stiglitz, Joseph. 1998 "More Instruments and Broader Goals· Moving towards the Post–Washington Consensus." WIDFR Annual Lectures 2 United Nations University World Institute for Development Economics Research, Helsinki, Finland Processed.

Summers, Robert, and Alan Heston 1991 "The Penn World Tables (Mark 5) An Expanded Set of International Comparisons, 1950–1988 " *Quarterly Journal of Economics* 8(May) 327–68.

Svensson, Jakob 1998 "Investment, Property Rights, and Political Instability Theory and Evidence " *European Economic Review* 42(7, July) 1317–41

Taylor, Charles, and Michael Hudson 1972 *World Handbook of Political and Social Indicators* New Haven, Conn · Yale University Press

Taylor, Charles, and David Jodice 1983 *World Handbook of Political and Social Indicators.* New Haven, Conn Yale University Press

——— 1988 *World Handbook of Political and Social Indicators* New Haven, Conn · Yale University Press.

Temple, Jonathan 1999 "The New Growth Evidence." *Journal of Economic Literature* 37(1) 112–56

Temple, Jonathan, and Paul Johnson 1998 "Social Capability and Economic Growth." *Quarterly Journal of Economics* 113(3) 965–90

Tornell, Aaron 1997 "Economic Growth and Decline with Endogenous Property Rights " *Journal of Economic Growth* 2(3) 219–50

World Bank 1993. *The East Asian Miracle Economic Growth and Public Policy.* New York Oxford University Press.

——— 1994 *Adjustment in Africa Reforms, Results, and the Road Ahead* New York: Oxford University Press.

——— 1997. *World Development Report 1997 The State in a Changing World* New York Oxford University Press.

Wright, Lindsay 1982 "A Comparative Survey of Economic Freedom " In Raymond Gastil, ed , *Freedom in the World.* Westport, Conn . Greenwood Press

World Bank Research Observer Cumulative Index, 1994–1999

Titles

Authors

International Approaches to Global Climate Change

Richard N. Cooper

This article surveys the issues involved in slowing the climate change induced by global emissions of greenhouse gases, especially carbon dioxide. It addresses the possible social and economic impacts of global warming, the elements involved in evaluating the pros and cons of steps to reduce those impacts, and the issues involved in engaging most of the world's states in a cooperative endeavor to reduce greenhouse gas emissions. It expresses doubts about the efficacy of a global approach based on national emission targets, such as those set by the 1997 Kyoto Protocol, and favors instead mutually agreed actions focused on a common emission tax. It also discusses issues of compliance with an international agreement to reduce emissions, actions states can take in the absence of international agreement, and contingency actions that might be considered if the problem proves to be more serious than now seems to be the case.

Human activity since 1800 has resulted in the emission of great volumes of gaseous materials into the atmosphere. Some of these gases—notably carbon dioxide, methane, and chlorofluorocarbons (CFCs)—absorb Earth's radiation, leading potentially to a warming of Earth's surface, which in turn could alter the world's climate. At the molecular level, CFCs are the most potent "greenhouse gases," but carbon dioxide has been emitted in greatest volume, largely from clearing forests and from burning coal and oil, and has the longest life in the atmosphere, thus accumulating over time. Atmospheric concentration of carbon dioxide in recent years has reached 360 parts per million (ppm), compared with about 280 ppm around 1800, and on some current projections is headed to 700 ppm (two and a half times preindustrial concentrations) by 2100 (Council of Economic Advisers 1998).

Wide scientific consensus suggests that under these conditions Earth's surface will become warmer on average, with temperature increases being higher in the higher latitudes. Sea levels will also rise, partly from melting glaciers but mainly from thermal expansion. Average global precipitation will increase.

The World Bank Research Observer, vol 15, no 2 (August 2000), pp 145–72

Beyond these general effects, the consensus dissipates. Earth's atmospheric phys ics and chemistry are complicated and not well understood; neither is the relation ship between the atmosphere and the oceans nor that between climate and the bio sphere (all forms of life). As a result, few agree on the rate at which chemical o biological processes take carbon dioxide out of the atmosphere, on the influence o greater warming and evaporation on cloud formation (which affects the extent t which the sun's rays are reflected away from Earth's surface), on the rate at which th oceans absorb heat from the atmosphere, and on a host of other relevant issues. Thu there is little agreement on either the ultimate extent or the pace of warming for an given trajectory of greenhouse gas emissions.[1]

Neither is the future trajectory of emissions known with high confidence, althoug continued "business-as-usual" economic growth can be expected to result in eve greater consumption of fossil fuels for many decades. The Intergovernmental Pane on Climate Change (IPCC 1996b) assumes in its main "business-as-usual" case global growth of 1.1 percent a year in emissions of carbon dioxide between 1990 an 2100. At this rate, and in the absence of actions to reduce the growth of greenhous gas emissions, the generally accepted range of warming over the course of the nex century is 1.0–3.5°C, with a best guess of perhaps 2.0°C, and an equivalently uncer tain rise in sea level of about half a meter.[2]

Confronted with these possibilities, the international community agreed at Rio d Janeiro in 1992 on a Framework Convention on Climate Change and at Kyoto i 1997 on a protocol that committed the countries listed in annex I of the framewor convention to reduce emissions by the year 2012 to estimated 1990 levels or up to percent lower, with the targets varying from country to country.[3] The countri listed in the framework convention include the 24 members of the Organisation fo Economic Co-operation and Development (OECD) in 1995 plus Central and East ern Europe and some successor states to the Soviet Union. The Kyoto Protocol ha not yet come into force, and indeed the Clinton administration indicated it woul not submit the document for Senate ratification until "significant participation" b leading developing countries was assured, as requested by Senate resolution.

Social and Economic Impacts

Several key potential economic effects of global warming have been identified Perhaps the most important is the effect on global food production. But concern have also been expressed about health, rising sea levels, and the general amenity o life. In addition, climate changes will have significant effects on nonhuman eco logical communities.

Agriculture

The impact of climate on agriculture depends on the detailed effects of climate change, particularly the regional and seasonal changes in precipitation. But the adverse effects (if any) that might occur under that heading must be assessed against the facts that plants rely on carbon dioxide as a major input to production; that increased atmospheric carbon dioxide, taken alone, would actually increase agricultural yields; and that agricultural producers around the world, but especially in temperate zones, have a demonstrated capacity for adapting to a variety of changes in their economic as well as physical environment. Comprehensive work on agricultural response to global climate change is still in an early stage, but several studies suggest that changes of the likely magnitude will not have a significant impact on global food output, although significant effects might be experienced at the regional level. For example, Schimmelpfennig and others (1996) show that 20–30 percent declines in output of grains at two locations in the United States under substantial increases in temperature (4–5°C) are greatly moderated, or even converted into increases, with plausible adaptations by farmers; the carbon dioxide fertilization effect would ensure increases for all three products—maize, soybeans, and winter wheat (Reilly 1998:246). Similarly, Darwin and others (1995) show that under four climate models, declines in global production of cereals caused by a doubling of atmospheric carbon dioxide become modest increases when farmer adaptation and market adjustments are allowed. As table 1 shows, small declines in nongrain food production are more than offset by increases in cereal and livestock (Meyer and others 1998).[4]

Although global food production does not seem to decline with global warming—on the contrary—the regional distribution is uneven. Production rises in the higher latitudes, partly because of an increase in arable land, and tends to fall in the

Table 1. *Percentage Changes in Global Production of Cereals by Climate Change Scenario*

World	*No adaptation*	*Land use fixed*	*No restrictions on use*
GFDL	–23.5	–0 6	0.3
GISS	–22.6	0 2	0 9
OSU	–18.6	–0 5	0 2
UKMO	–29 3	–0 2	1.2

Note "No adaptation" represents the additional quantities firms would be willing to sell at 1990 prices under the alternative climate The next two columns represent changes in equilibrium quantities, under new equilibrium prices, with land use fixed and with no restrictions on land use The results are based on equilibrium scenarios in four climate models that assume a doubling of atmospheric carbon dioxide those developed at the Geophysical Fluids Dynamics Laboratory (GFDL), the Goddard Institute of Space Studies (GISS), Oregon State University (OSU), and the United Kingdom Meteorological Office (UKMO)

Source Adapted from Darwin and others (1995), Reilly (1998)

climate change, since it will be difficult to persuade people to undertake costly or painful actions for benefits that accrue mainly to others.

Nonhuman Ecological Systems

Although human beings have demonstrated a remarkable capacity for adaptation to a variety of conditions and new developments, the same cannot be said about other species. Rapid climate changes may find many species unable to adapt in the time required.

Rising sea levels will affect natural ecosystems, particularly wetlands, which are known for their high levels of biological activity. Moreover, some human adaptation may come at the further expense of wetlands. For example, Nicholls and others (1995) estimate that without countervailing measures, rising sea levels will adversely affect 56 percent of the world's wetlands, and this figure rises to 59 percent when allowance is made for measures to protect human settlement. Existing wetlands are in rapid decline, however, for reasons that have nothing to do with climate change. If wetlands are to be preserved, affirmative human action will have to be taken, whether or not climate change threatens them.

Warming of temperate and northern latitudes will alter the natural vegetation, which in turn will alter the natural fauna. But trees take a long time to grow, and species move in nature only as rapidly as seeds can be carried by wind or creatures into newly habitable territory. Humans can assist other species to adapt to the new conditions, however, provided the requisite knowledge is available and the issue is considered sufficiently important.

Speciation is much higher in the Tropics than in higher latitudes. Microecosystems flourish in that climate, where highly specialized plants and (especially) animals with limited ranges have developed. Fortunately, temperature increases are likely to be least in tropical zones, but changes in precipitation and carbon dioxide fertilization will permit some species to flourish at the expense of others, possibly driving some to extinction.

A Framework for Collective Decisionmaking

Concerns about global climate change have led to pleas and indeed to some national commitments to slow or reverse the growth of greenhouse gas emissions. It is useful to identify the structural characteristics involved in attempting to mitigate global warming through formal collective action. There are three key features.

First, climate change brought about through an increased atmospheric concentration of greenhouse gases is a global issue, since whatever their earthly origin, the gases are widely dispersed in the upper atmosphere. Effective restraint must there-

fore involve all (actual and prospective) major emitters of greenhouse gases. Rich industrial countries account for most of the emissions today, but the Soviet Union was a major contributor before its dissolution and economic collapse in 1991, and the area can be expected to become a major source with economic recovery. Rapidly growing developing countries will become major contributors within a time frame that is relevant for managing the issue. Thus while the same requirements need not be imposed on all countries from the beginning, the agreement needs to be structured so that all countries will eventually participate. By one estimate, for example, full implementation of the Kyoto Protocol and continuation at the prescribed lower emission levels by annex I countries would slow the increase in average global surface temperature in 2050 by only 0.05°C, from an increase of 1.4° to 1.35°.

Second, the rewards from restraints on greenhouse gas emissions will come in the (politically) distant future, while the costs will occur in the political present. Moreover, the rewards are highly uncertain. The residents of some of today's countries, such as Canada, Russia, and perhaps the United States, may even expect to benefit from moderate climate change. It will thus be difficult to persuade people that they should make sacrifices in their own living standards for the sake of uncertain gains to their grandchildren and great-grandchildren, and to the grandchildren of others, remote in distance. The wide distribution of expected but distant benefits in response to collective action today provides an incentive for every country to encourage all to act but then to avoid acting itself—the so-called free-rider problem.

Third, the pervasiveness of the sources of greenhouse gas emissions—notably fossil fuel use, rice cultivation, and cattle production—implies that restraint will involve changes in behavior by hundreds of millions—if not billions—of people, and not merely by 180 or fewer governments, as in the typical treaty. Thus the most important part of an effective regime to limit climate change involves not the relationships among states but the effective influence of governments on the behavior of their citizens.

No major legally binding regulatory treaty touches all of these characteristics to the same degree. Typically, treaties apply to the actions of either governments themselves or a relatively few firms in a relatively few countries, as in the cases of halting nuclear testing or limiting production of CFCs. The Convention on International Trade in Endangered Species perhaps comes closest in its comprehensiveness; it requires states to prohibit international trade in an agreed list of products. The Chemical Warfare Convention is extremely intrusive in its monitoring requirements but has not yet come into force. The various agreements for management of international fisheries require cooperation from hundreds of fishermen, but with a few exceptions, these pacts have not been notably successful.

These three structural factors make collective decisions regarding actions to mitigate global climate change exceptionally difficult. Serious mitigation necessarily involves major reductions in the actual and prospective consumption of energy based

on fossil fuels (especially coal-fired electricity generation and the use of oil products for heat and transport). Because such consumption is at the very heart of modern industrialized economies, the costs of mitigation are both the economic and the psychological adjustments that must be made to move away from current energy systems and from wet rice and cattle production, which, along with leaks from gas and oil refining and distribution systems, are the main sources of methane from human activities. Moreover, the likelihood that the distribution of costs and benefits will be highly uneven across nations further complicates the task of reaching international agreement.

It is natural for an economist to compare the overall benefits of any proposed change with the overall costs required to make the change. Many noneconomists reject cost-benefit analysis as an artifact of calculators who ignore or underrate basic human values. But this rejection is simply an intellectual mistake; everyone who urges a change in policy (or resists one) is at least implicitly comparing costs with benefits. The disagreement, rather, is about how best to measure the alleged benefits and costs of the proposed change. Thus when Krause, Bach, and Koomey (1992) argue that on no account should the average global temperature be allowed to rise more than 2.5°C (the upper limit of Earth's temperature in the past 2 million years), implying, according to their worst-plausible-case calculations, that no more than 300 billion tons of additional carbon can be emitted into the atmosphere, they are implicitly arguing that the benefits of severe mitigation action are infinitely great and warrant any finite cost to achieve them. The authors are expressing an extreme degree of aversion to environmental risk. Others may properly disagree with such extreme valuation.

Table 3 illustrates the range of marginal benefits per ton of carbon emissions avoided and the range of marginal costs of reducing emissions by a ton of carbon for two different targets: a return to 1990 emissions and a 20 percent reduction from 1990 emissions. All the estimates must be taken as illustrative, as the methodologies used are quite different and are incomplete. With this caveat, these estimates show that costs generally exceed benefits, especially for 20 percent reductions in emissions. The wide variation suggests that the methodology for estimating costs and benefits can stand considerable improvement. Not surprisingly, the costs of reducing emissions rise with the magnitude of the reductions. Kram (1998) and Stavins (1999) suggest that the increases can be very steep.

The Discount Rate

Cutting greenhouse gas emissions involves incurring costs long before the benefits are registered. A comparison of near-term costs with future benefits requires use of a discount rate (or stream of rates) to put both benefits and costs into present value. Much has been written about the appropriate choice of a discount rate and the prin-

Table 3. *Selected Estimates of Benefits and Costs for Global Marginal Abatement of Carbon Dioxide*
(U S dollars per ton of carbon)

Benefit study	*Marginal benefit*[a]	*Cost model*[b]	*Marginal cost*	
			Stabilization	*20 percent reduction*
Ayres and Walter	30–35	Jorgenson-Wilcoxen	20	50
Nordhaus	7	Edmonds-Reilly	70	160
Cline	8–154	Manne-Richels	110	240
Peck and Teisberg	12–14	Martin-Burniaux	80	170
Fankhauser	23	Rutherford	150	260
Maddison	8	Cohan-Scheraga	120	330

a For most studies the marginal benefit increases over time The estimates presented here correspond to the period 2001–10.

b Cost estimates are from a study by the Energy Model Forum of Stanford University, which ran 14 different cost models using common assumptions and standardizing for the emission reduction scenarios shown above

Source IPCC 1996a (tables 6 11, 9 4).

ciples that should undergird the choice. Theoretical and some practical economists have been fascinated by the Ramsey model of savings, which suggests that the optimum social rate of time preference (r) can be expressed by the simple equation $r = \tilde{n} + \grave{e}g$, where $\tilde{n}$ is the pure rate of time preference, $\grave{e}$ is the elasticity of marginal utility with respect to additional consumption, and g is the growth rate of per capita consumption (see Nordhaus 1994 or IPCC 1996a, ch. 4, for an explanation). Plausible numbers for these variables lead to discount rates ranging from 0.5 to 3 percent (IPCC 1996a; Cline 1998).

The underlying rationale for avoiding or mitigating climate change is to benefit future generations. Yet to undertake investments in the near future that yield, say, 2 percent over the next century does a great disservice to future generations compared with other investments that we have strong reason to believe yield much higher returns many years into the future, if not for an entire century. Surely, in the interest of future generations, society should prefer high-return investments to low ones.

There is evidence that returns to education in developing countries exceed 20 percent (Psacharopoulos 1985, 1994). Returns to college education for a male in the United States reportedly equal 13 percent (Council of Economic Advisers 1996). A study of more than 1,000 projects completed by the World Bank in the 1970s and 1980s yielded an average (prospective) return of 16 percent (Pohl and Mihaljek 1989). The World Bank and the U.S. government have stated threshold returns of 10 percent for evaluating prospective investments (recently reduced to 7 percent by the U.S. government). The corporate sector of the U.S. economy, one that is relatively rich in capital by global standards, yields an average pretax real return well above 10 percent. For all these reasons, 10 percent seems to be a reasonable rate of discount. A

high discount rate, of course, gives less weight to benefits (and costs) in the distant future. But that implication alone is not sufficient reason to reject it.

Maurice Scott of Oxford University has suggested 4 percent (reported in Beckerman 1996), partly because that has been the real yield on low-risk government bonds in recent decades. But even if resources can be extracted from the public at 4 percent, they should be invested in those activities with high (social) return. Only after we exhaust 10 and 7 and 5 percent opportunities should we accept investments with prospective yields of only 4 percent. Otherwise we deprive either future generations or our own generation unnecessarily.[5]

Some observers object to citing data on observed rates of return on grounds that actual decisions made today and in the past were not made under ideal conditions but instead reflect a number of imperfections both in markets and in the processes for making collective decisions. It would take me too far afield to explore this contention in relevant detail. Let me just stipulate that the real world is messy and that actual decisions (and market outcomes) deviate from any given set of ideal standards The same observation applies to actions to mitigate greenhouse gas emissions A plausible argument must be made that allowance for the various imperfections will raise the after-the-fact returns to mitigation actions relative to the observed returns on other investments.

The debate over the choice of a discount rate can be interpreted as an effort to reduce or eliminate imperfections in collective decisionmaking on public expenditures in general. But if such imperfections are important, and if other public investments seem to leave future generations still better off, low discount rates should also apply to those higher-yield investments. Advocates who apply them only to climate change must do so either because they believe the political prospects are better for improving collective decisionmaking on global climate change than for other, higher-yield public investments or because they must prefer mitigation of greenhouse gas emissions on some unstated grounds, not captured in the usual reckoning of costs and benefits over time, and want to support such actions—or both. In any event, it would be useful and desirable to open these considerations explicitly to wider discussion.

This sounds like common sense. What can be the objection to it? One possibility is that while the return on investment A (education, say) exceeds that on investment B (mitigation) in the near term, the reverse is true in the long run because of a secular decline in returns to investment A. Normally one could switch to B investments as returns to A drop below those to B. Investing in B now would be preferable only if for some reason it would be too late to switch to B investments later, after returns to new A investments fell.

This type of configuration is theoretically possible, but a plausible case must be made for both parts before one can conclude that A should be rejected in favor of B at the outset. In the standard neoclassical economic model, the returns to capital are

assumed to fall steadily as the ratio of capital to labor (and other factors) rises. But in historical—as distinguished from analytical—time, technical change has constantly increased the returns to (new) capital, and there is no reason to believe that the process will stop during the next century. Thus if returns to class A are high now relative to B, they are likely to remain so.

Distributional Considerations

The IPCC (1996a, ch. 4) authors seem to reject the efficiency argument that is emphasized here, not on the foregoing grounds but on the basis of equity. In essence, they argue that one cannot ethically say that investment A is superior to investment B even if it yields higher total future benefits if those who experience losses as a result of the investment are not actually compensated (in the absence of a social welfare function that indicates the relative weights that should attach to winners and losers).

This is a valid theoretical point. But if taken literally and applied seriously, it is a prescription for total inaction, especially when time frames as long as 100 years are under consideration. First, there is no collectively agreed social welfare function, and no prospect of one at a global level, so we cannot generally weigh winners against losers, especially over so long a time. Second, we cannot possibly know distant future winners and losers from our actions today. Try, for example, to identify the winners and losers from completion of the U.S. transcontinental railroad in 1869, or to forecast the winners and losers 100 years hence from construction of the Three Gorges Dam in China. And third, even if we had the requisite knowledge about future winners and losers and our preferences among them, we cannot bind future generations to adhere to those preferred outcomes. If we make rules, future generations can unmake them. If we plant trees, they can cut them down. If we consume less coal, they can consume more—and may actually do so because it is more readily available to them. The one legacy that cannot be reversed (short of a collapse of civilization) is enhanced knowledge—both a deeper understanding of nature and improved technology.

We should be concerned above all with passing more knowledge and higher incomes to the next generation than we received from our parents and should allow its members to decide how to distribute them. They will do so in any case, regardless of what we think. This is not to suggest that we should be completely indifferent to distributional effects. Our actions will affect the initial distribution of the next generation, and collectively we may want to avoid certain actions on grounds that we do not like their distributional effects. But here I mean the direct consequences of our actions, on which it may be possible to get collective agreement to avoid imposing extreme losses on certain classes of people. We cannot carry this logic into the more distant future, however, for the reasons already mentioned: we cannot possibly know the future impact on people (for one thing, we do not even know where they will be), and we cannot commit future generations to our preferences even if we did.

In any case, it is rather odd to urge costly action now for the sake of poor people in the distant future when we are not willing to take very costly action now for the sake of reducing poverty today. We have actual evidence on the amounts we are willing to spend, individually and collectively, in the name of economic development of today's poor countries: about 0.3 percent of the gross domestic product (GDP) of the rich countries, and only a portion of that is devoted to reducing poverty as such. If we are really concerned about the impact of possible future climate change on poor people, we should take more active steps to reduce their current poverty. That would improve their capacity to adapt to such climate changes as may take place and to take mitigation actions themselves.

If there is a general disposition within the rich countries to help people in poor countries, the best way to do it is probably through education. Education has at least three advantages with respect to mitigation of climate change. First, the rate of return seems to be substantially higher, at least on the estimates that have been made so far. Second, it is harder for future generations to undo the redistribution favored by this generation, since educated parents are likely to want to see their children educated. Third, education increases the capacity of any society, and of individuals, to adapt to changing circumstances, including but not limited to changes in climate

Risk Aversion

It is widely taken for granted, at least on big issues, that people dislike uncertainty, they have an aversion to risk and are willing to pay something to reduce risk. This is the attitude that underlies the willingness of individuals to take out fire or liability insurance, to mitigate the possible costs of uncertain and perhaps even improbable unfavorable events.

The uncertainties associated with mitigating global climate change and its attendant costs are at least as great—and probably greater—than the uncertainties associated with other investments that could be made today; given risk aversion, one might thus conclude that costly mitigation actions should not be undertaken. However, the payoff from mitigation actions now will be greatest if the magnitude of global climate change and the associated costs turn out to be high, even if that is judged to be a low probability. Of course, if the costs associated with global climate change are low, any investment in mitigation actions will have a low or negligible return. But such investment may still be worthwhile as insurance against an uncertain but possibly costly contingency.

How do these considerations influence the discount rate? The precise answer is not at all straightforward, unless the uncertainty itself is related in a particular way to the passage of time. Roughly speaking, however, where an uncertain outcome (the future payoff from mitigation actions) is negatively correlated with the overall economic prospects and where the uncertainty grows exponentially with time, some

deduction from the discount rate used to evaluate mitigation actions is warranted. How much? That depends in detail on the nature of the uncertainty, an issue that needs much greater discussion, and on the degree of aversion to risk. But presumably it was this sort of consideration that led U.S. policymakers in 1980 to stipulate a discount rate of only 7 percent for publicly financed energy-related projects, 3 percentage points lower than the general standard for government investments. Serious disturbances in the field of energy, unlike other areas, can lower gross national product by a multiple, so some component of the energy investment can be regarded as an insurance premium designed to attenuate the economic impact of large disturbances in the world oil market.

Nordhaus (1994) models the determinants of climate change and the effects on emissions, temperature increase, warming damage, world output, and so on, focusing on optimal mitigation policies. He recalculates the optimal mitigation policy, taking into account uncertainties in those factors that influence climate change. Not surprisingly, the optimal reduction in greenhouse gas emissions, and the carbon tax required to achieve it, are higher in the presence of these uncertainties than they would be with confident best-guess projections. Concretely, the optimal carbon tax during the 1990s under the uncertain conditions postulated by Nordhaus is $12 a ton, compared with less than $5 a ton on the best-guess projection. Of course, the optimal policy is likely to change in response to new knowledge.

What about the possibility of truly disastrous outcomes as a result of global warming? Although the scientific community does not put a high probability on any of them, three are sometimes mentioned: first, warming sufficient to release the extensive methane contained in the Arctic permafrost, leading to a strong and possibly rapid reinforcement of warming; second, warming sufficient to break up the Antarctic ice dam and raise the oceans several meters, rather than half a meter, higher; and third, glacial melting in Greenland of a volume and character sufficient to deflect southward the warm North Atlantic currents, paradoxically making Europe a much *colder* place.

These possibilities, however remote, raise the question of risk aversion and how much insurance societies are willing to buy against improbable but highly costly contingencies. There is no doubt that individuals vary greatly in their degree of risk aversion and that commercial insurance policies do only a modest job of bringing these diverse preferences into harmony at the margin. The market for differences in preference regarding risk is much less well developed than the market to take advantage of differences in time preference. Each society has its own mechanism, through the political process, for deciding and acting on the degree of collective risk aversion. But the mechanism for the world as a whole is much less well developed, being mediated through diplomatic conferences such as those at Rio in 1992 and Kyoto in 1997, followed by public debate and ratification.

The political process, while essential for making decisions on collective risk, contains some serious weaknesses. Most notably, the discussion is not conducive

to honesty and straightforwardness. Some risk-averse parties will exaggerate the risks to persuade those who are less risk-averse than themselves. Others will attempt to minimize the estimated costs of early action or suggest that they can be borne by nonvoters (for instance, corporations). And some will cite legitimate concerns about greenhouse warming to encourage society as a whole to adopt a certain "lifestyle," such as relying less on the automobile, a technology that has been liberating for many people. By the same token, those who expect to bear the costs of political decisions in response to concerns about climate change will tend to minimize the risks and exaggerate the costs of mitigation—or even deny that a problem exists. In short, we should be on guard against strong but misleading or exaggerated arguments by all sides.

One way to deal with a potentially important problem that is subject to profound uncertainties is to establish a framework for action with broad participation and institutional procedures for integrating new information into decisions, as Schmalensee (1998) has urged, but to avoid costly actions that may turn out to be mistaken

International Burden-Sharing

Suppose that the international community decides that steps should be taken to reduce greenhouse gas emissions and that many or all nations should be involved What might be included in such a treaty? One approach, reflected in the Kyoto Protocol, is to impose national targets on emissions, possibly permitting some of the allowed emissions to be transferred from one nation to another. A second approach, which has received less attention, would call for a set of actions that states would agree to undertake. In my view, mutually agreed actions have better prospects of mitigating emissions than national targets.

Setting National Targets

When quantitative targets are imposed within countries, they almost universally respect recent history, being set at levels roughly in proportion to recent use (for instance, oil refinery throughput or emissions of sulfur or harvests of halibut). Targets based on emissions in a fixed base year such as 1990, as at Kyoto, have a similar character. They in effect allocate property rights to the existing tenants, accepting the right of ownership by virtue of possession or use. Targets allocated on this basis will be completely unacceptable, however, to countries that are or expect to be industrializing rapidly and that anticipate a disproportionately rapid growth in demand for fossil fuels. They will argue that most of the existing stock of greenhouse gases was emitted by rich countries and that those countries should therefore bear a disproportionate responsibility for cutting back. Thus developing countries did not

commit themselves to reduce emissions at Rio or Kyoto; within Europe, Greece and Spain expressed similar reservations.

Some observers have suggested that simple distributive justice would require that emission rights be based on population. Such an allocation would favor heavily populated poor countries such as Bangladesh, China, India, Indonesia, and Nigeria. To be meaningful in limiting climate change, such allocations would require drastic cutbacks in emissions by rich countries, implying radical reductions in living conditions there if implemented quickly. Targets based on population would be insensitive to varying resource endowments (such as hydroelectric power) and to the fact that countries depend on vastly different fuel mixes and have different levels of fuel consumption.

Reductions in living standards could be mitigated, but not avoided, by the sale of unused emission rights from poor to rich countries (discussed further below). But the financial transfers involved if emission rights were based on population would be immense relative to foreign assistance today—far greater than is likely to be politically tolerable. If carbon emissions were to take a plausible value of $100 a ton, for instance, the typical American family of four would have to pay $2,200 a year to sustain its current (direct and indirect) average level of emissions of about 26 tons a year, 22 tons over its per capita allocation (roughly 6 billion tons of carbon emissions a year divided by a world population of roughly 6 billion people). Total U.S. transfers to the rest of the world would amount to $130 billion a year, more than 10 times current U.S. foreign aid expenditures. Moreover, the transfers in practice would be made to governments, despite the underlying moral rationale for basing targets on population, and many people would question the desirability of transferring large sums to governments whose responsiveness to the needs of their own citizens has been indifferent or worse.

A natural compromise has been suggested: base the national targets on GDP (or recent past emissions) initially and gradually convert them to population-based targets over, say, 25 years. Here, however, we encounter some unpleasant arithmetic with respect to population-based emission rights. In 1995 India's per capita income (on a purchasing power basis) was about 5.2 percent that in the United States. Suppose that per capita income in India grows at 5 percent a year over the next 25 years and per capita income in the United States grows at 1 percent a year (this is a plausible scenario, although in reality the gap in growth rates is not likely to be so wide). Under those assumptions, Indian per capita income 25 years from 1995 (in 2020) would still equal only 14 percent of per capita income in the United States, and per capita consumption of energy would be many times higher in the United States than in India. Thus, after 25 years, one of three possible scenarios would emerge: India would not be effectively constrained; the United States would be very tightly constrained; or (under tradable emission permits) there would be huge transfers from the United States to India. The sense of global community is not likely to be great

enough by 2020 to sustain such large transfers—it is not that great *within* the United States today—and in any case such large transfers, either to governments or directly to citizens, would probably not be desirable, as some of the highly oil-dependent countries have discovered.

Perhaps the most reasonable way to allocate emission rights and the obligation to reduce emissions would be to calculate a "business-as-usual" trajectory of emissions for each country on the basis of recent history, development prospects, and past experience with the evolution of greenhouse gas emissions in relation to economic development. Then each country could be charged with reducing emissions by a uniform percentage, chosen in relation to global reduction requirements, relative to the assigned trajectory. Of course, even if this principle of allocation of rights and responsibilities were accepted as reasonable, the debate would simply shift to the choice of trajectories for each country. Developing economies aspire to grow rapidly. The Republic of Korea and Taiwan, China, have demonstrated that growth of more than 8 percent a year for three decades is possible. Most developing countries will set their aims similarly high and insist on energy-consumption growth to support them. They will be reluctant to accept lower emission targets without assurance that the technology will be available to achieve their growth targets with the lower emissions. Who is to say they are wrong?

Implementing the Agreement

Once national targets have been established, they must be translated into conditions that induce firms and households to change their consumption patterns to avoid certain activities. For large firms—generators of electricity, say—that could perhaps be done by fiat simply by setting quantitative limits for each generating plant. But for most economic agents the only practical way to alter behavior is to create price disincentives; that is, to tax the activities that generate the emissions.

Every international agreement must address the question of compliance and the associated question of monitoring behavior to discover if it deviates from the treaty requirements. In principle, given the objective, all significant greenhouse gases should be covered. In practice, given the many actors involved and the many sources of emissions, such broad coverage would be impossible to monitor and police. For practical reasons, therefore, attention is usually focused on fossil fuel consumption (plus a few other concentrated emitting activities, such as cement production). Monitoring such consumption is more or less manageable, since most of it must pass through some relatively narrow choke points (gas pipelines, oil refineries, electricity generating stations). Most coal production can be monitored at the mine-head or on the barges and railroads that transport the coal.

But this still leaves out a lot of emissions. Only about half of the greenhouse gas emissions (measured by radiative forcing, which is what is relevant for climate change)

since 1850 have come from the burning of fossil fuels (IPCC 1996b). The rest have resulted from the burning of tropical forests, cultivation of livestock and rice, dumping of garbage, and losses from gas pipelines. Omitting these sources would misrepresent total emissions. The Kyoto Protocol covers 24 gases, including methane and nitrous oxide, in addition to carbon dioxide (Council of Economic Advisers 1998). Monitoring emissions of all these gases will be difficult—and probably impossible—if developing countries are covered by the requirement.

If the fossil fuel carbon emission targets for rich countries are so demanding, how are they to be met? Conceptually, there are four ways: more efficient conversion of fossil fuels to usable energy in existing plants; switching away from fuels that are high in carbon per unit of energy produced (basically, shifting from coal to natural gas); building new plants that use less carbon per unit of usable energy (for example, nuclear power plants); and reducing end-user demand for energy. Unfortunately, the scope is limited for further change at the easiest monitoring points. Obsolete generating plants can be replaced with more efficient or less carbon dependent ones, but replacement demand in the OECD countries will be modest during the next 20 years, and replacing generating plants before they become obsolete is extremely expensive. In developing countries the demand for electric power is rising rapidly, so most of the generating capacity that will be available in those countries in 2010 could in principle be designed to use technology with reduced carbon dependence.

The consequence is that most of the reduction in the rich countries must come from downstream, at or near the points of final demand, where the number of consumers is greatest. Quantitative rationing is neither desirable nor feasible in market economies, so the reductions must be achieved by some combination of price (dis)incentives, exhortation through publicity, and education on best practice. Many consumers are not aware of the ways they can conserve energy without making radical changes in lifestyle. But in any case the key to success is not at the intergovernmental treaty level but rather in the incentives each government can provide to its own citizens. A treaty merely provides a vehicle for rough "burden-sharing" across countries and some international discipline in pursuit of the targets.

Opportunities for reducing emissions in new electric generating plants and other new industrial facilities will be greater, and the marginal cost lower, in developing countries than in mature economies. That realization has led to an emphasis on "joint implementation," a procedure whereby agents in rich countries can obtain credit against national targets in their own countries for making emission-reducing investments in developing countries. The idea is attractive. But under the Kyoto Protocol, the developing countries do not have national targets. Therefore, avoiding reductions in emissions in rich countries by investing in poor countries by itself will not reduce global emissions, since much investment must and will be undertaken in developing countries anyway. Reducing global emissions can be accomplished only by establishing detailed criteria for "additionality" in emission-reducing new invest-

ments—that is, by establishing norms by country and by project for least-cost power generating or energy-using investments and then counting reductions in emissions relative to such norms. The norms themselves would be changed as technology advanced. Establishing such norms would be both complicated and controversial because it would necessarily involve both judgment and approximation.

Imposing a Carbon Tax

There is an important alternative to setting national emission targets: an international agreement on a set of actions calibrated to achieve the desired emissions. Because to accomplish their quantitative objectives governments must in any case create the appropriate behavior-altering incentives for their citizens, and because setting a national allocation of global emission rights for both rich and poor countries is likely to prove so contentious as to be impossible, it may be easier simply to agree on a common use of instruments. For problems such as reducing emissions, the favorite instrument of economists is to tax the offending activity. All countries would agree to impose a common carbon tax, which would increase the price of fossil fuels in proportion to their carbon content. Such a tax would have at least two major advantages. First, it would encourage reduction of emissions where that can be done at least cost. All emitters would have the same incentive to act, but only those who saved more in tax payments than it cost to reduce emissions would undertake reductions; others would simply pay the tax. A carbon tax would encourage users everywhere to switch to natural gas (which would benefit Iran, Russia, and other countries with large gas reserves) and, more important, would prompt consumers generally to conserve fossil fuels. Second, a carbon tax would generate funds for governments that have trouble finding sources of revenue that do not have negative effects on economic incentives to work, save, or undertake commercial risks.

A common carbon tax would be easy to monitor. Enforcement would be more difficult, but all large countries except Cuba and the Democratic Republic of Korea hold annual consultations with the International Monetary Fund (IMF) on their macroeconomic policies, including the overall level and composition of their tax revenues, and the IMF could provide reports to the agency that monitors the treaty governing greenhouse gas emissions. Such reports could, if necessary, be supplemented by international inspection of the major taxpayers (for instance electric utilities) and of the tax agencies of participating countries.

Such a regime would present a major problem to democratic countries, however, because taxation goes to the heart of parliamentary prerogative, and most democracies will not welcome taxation by international agreement. Moreover, even modest energy taxes are politically unpopular.

Two additional problems need to be mentioned, neither insuperable. The first is that countries tax energy (especially oil) differently, and some countries continue to

price both coal and oil well below world levels. Should a uniform tax be levied on an uneven initial condition? If existing pricing practices are taken to reflect national preferences for allocating resources, a case can be made that the new carbon tax should be uniform, regardless of the initial tax burden. Of course, national policies would have to be monitored to ensure that the effect of the new tax was not undermined by other changes in tax or subsidy policy. Alternatively, the treaty could simply require a minimum national tax on emissions from fossil fuels, allowing existing taxes to drop toward that minimum, as advocated by Nordhaus.

The second problem concerns the disposition of revenue. Available estimates suggest that to have a significant impact on emissions, the tax might have to be substantial and thus would generate a great deal of revenue. To whom should this revenue accrue? Oil-producing states will suggest that if oil is to be taxed, they should levy it and get the revenue—indeed, that is what the attempts by the Organization of Petroleum Exporting Counties to control oil prices amount to. Oil-consuming countries, however, would feel doubly aggrieved if they charged more for oil to discourage its consumption and yet did not get the tax revenue; they would insist that the tax be levied on consumption and that the revenues accrue to them, not least so that they could reduce other taxes to ensure their continued prosperity and growth. In practice, the latter view is likely to prevail.

There is, however, a third possible claimant for the revenue: the international community. The international community has accepted a number of collective obligations that are cumulatively expensive. The most apparent are caring for refugees and peacekeeping operations, each of which cost the United Nations about $1.3 billion in 1995. Special assessments are now made for these activities, and several countries, including Russia and the United States, are in arrears. The regular UN budget runs $1.2 billion a year. In addition, donor countries finance the United Nations Development Programme and the International Development Association (at about $5 billion a year) to provide economic assistance to the poorest developing countries. Under the Rio framework convention, cooperation by developing countries in reducing emissions is conditioned on new financial support from the rich countries. Some or all of these activities could be financed in part by the tax revenues levied in pursuit of reduced emissions; obviously the major emitters, currently the rich countries, would pay most of the tax. But as poor countries develop, their contribution would increase automatically.

Estimates from several global energy-environment models suggest that a uniform reduction in carbon emissions from a "business-as-usual" baseline for each country or region would require very different carbon tax rates if that were the policy instrument used to reduce emissions. A uniform tax rate across the regions studied would result in quite different reductions from the baseline—as one would expect from the observation that countries around the world use energy with very different degrees of efficiency. Table 4 reports the per-ton carbon tax (in 1990 U.S. dollars) that would

Table 4. *Per-Ton Carbon Tax Required by 2050 to Achieve a 2 Percent Annual Reduction in Carbon Emissions from Baseline*
(1990 U S dollars)

Region	*Edmonds/Reilly*	*Manne/Richels*	*Green*	*Carbon rights trade model*
United States	1,096	208	340	754
Other OECD members	734	208	299	365
Former Soviet Union	325	990	180	2,245
China	341	240	67	1,109
Rest of world	1,012	727	329	763

Source: Dean and Hoeller (1993.153)

be required in five regions in the year 2050 to reduce carbon emissions by 2 percent a year from the baseline trajectory. Since the baseline trajectories project an increase in energy-related carbon emissions from roughly 6 billion tons a year in 1990 to 11 billion–19 billion tons in 2050, the 2 percent a year reduction would leave emissions that ranged between 3.3 billion and 5.7 billion tons in 2050, less than the 1990 levels.

By 2050 the world price (in 1990 dollars) of oil, 56 percent carbon by weight, is assumed to be $50 a barrel, two and half times its price in 1997. More recent projections suggest lower oil prices in the future, resulting from substantial improvements in the technology for discovering and extracting oil. Lower oil prices would lead to greater carbon emissions and thus to the need for greater taxes to bring them down to a target level. The price of coal, 75 percent carbon by weight, is assumed to be $60 a ton, about twice the recent price at points of exportation. Thus a tax in 2050 of $208 per ton of carbon would represent a 31 percent tax on oil and a 260 percent tax on coal. The loss in GDP engendered by this emission reduction program ranges (across the studies) from 1.3 to 4.9 percent in 2050 for the United States, from 2.3 to 6.4 percent for the countries of the former Soviet Union, and from 2.1 to 5.1 percent for the rest of the world—today's developing countries, minus China (Dean and Hoeller 1993). These results must be regarded as merely exploratory rather than definitive, but even the low estimates suggest a substantial cost to bringing energy-related carbon dioxide emissions below 1990 levels. The revenue these taxes would raise is also substantial. For instance, a carbon tax of $208 a ton in the United States would raise nearly $300 billion in revenue, 1.8 percent of estimated GDP in 2050. A carbon tax of $329 a ton in the rest of the world would raise $610 billion in 2050, nearly 3.2 percent of rest-of-world GDP in that year.

Trading Emission Rights

A gain in efficiency of emission reduction similar to that which would be achieved by a uniform world carbon tax can be achieved by allocating national targets to the

major emitters of carbon dioxide and allowing them to purchase or sell emission permits. A world market would quickly develop in such tradable emission permits, with a uniform world price. An emitter that could reduce emissions at a cost lower than the permit price would have an incentive to do so and would then sell its unneeded permits into the world market. An emitter that could reduce emissions only at a cost above the permit price could save money by buying enough permits to cover its excess emissions. Table 4 suggests that there would be much scope and mutual gain from a global market in permits, since the estimated costs of reducing emissions vary greatly from region to region in all the models.

The U.S. government, which has estimated the gains from trading emission rights, projects that the marginal cost for meeting the Kyoto target of 93 percent of 1990 emissions by 2012 would be about $200 a ton of carbon (calculated from Council of Economic Advisers 1998). If emission permits were allocated and traded, this cost would be reduced by 72 percent, to $56 a ton Americans would not meet the 7 percent reduction themselves but rather would buy permits, mainly from Russia, which would have an easier time meeting its Kyoto target of no change from 1990 because of the collapse of heavy industry after 1990 and because of the considerable scope that country has for improving energy efficiency. Adding some key developing countries such as China and India to the trading regime would reduce the price to an estimated $23 a ton, allowing Americans to buy more permits more cheaply. The developing countries, by the same token, would receive substantial payments for their surplus emission rights.

An effective market in emission rights requires that the trading parties have well-defined property rights. These parties could in principle be governments, but in most countries permits would have to be allocated to the relevant firms. The allocation of emission rights would be a nontrivial political issue because of the distributional implications. As noted above, the historical tendency is to allocate quotas on the basis of historical performance. But the emission rights would have substantial value, and there would be little social merit in allowing the grandchildren of today's emitters to continue to own the rights 50 years from now. The distortions over time from "grandfathering" the emission rights could be avoided by auctioning the rights to the emitters from the start, with the revenues accruing to governments. Some of the revenues could temporarily subsidize today's emitters, which would have the major burden of adjustment in the new regime.

Because of the need to establish clear property rights to emissions, such a regime could include only annex I countries under the Kyoto Protocol; other countries have no emission targets. The inclusion of key developing countries would require them to agree to (necessarily growing) national targets and to a mechanism for allocating national targets to the emitting firms. Such allocation would of course be a strong temptation to graft and corruption because the rights would have substantial value under current estimates.

A permit trading regime would, moreover, require careful monitoring and enforcement to ensure that parties that had sold emission rights in fact cut their emissions to the levels stipulated. Another issue would be the potentially large transfers of wealth from permit-buying countries to permit-selling countries, with the magnitude of the transfer depending not only on the price of the permits but also on the initial allocation of emission targets, a matter touched on earlier.

Of course, implementation of the Kyoto targets without extension or global trading would also have distributional implications across countries, brought about both through the relocation of high-energy-using activities (and the associated investment) to countries without targets and through changes in the terms of trade that would result from implementing the Kyoto regime and making the secondary adjustments to that implementation.

The Kyoto Protocol, if taken seriously, will be costly to implement, will have a sharp impact, and will be of little benefit to the climate. Signatory governments have not leveled with their citizens on the full implications of implementing the agreement. When those implications are known, the public is likely to balk; it is plausible that most annex I countries will not meet the ambitious targets of the Kyoto Protocol by 2012. A cynic might argue that politicians understand this fully and have undertaken a classic straddle to have their cake and eat it too: cater to the single-minded constituents demanding immediate action on climate change, while committing themselves to an international framework that is likely to prove unworkable—not exactly a tragedy, since it may be neither necessary nor desirable.

Of course, any regime that taxes fossil fuels (directly or indirectly), that excludes some countries from the control regime, and that permits (or prohibits) trades in emission permits will have substantial redistributive effects among countries. Energy-intensive industries will move to countries that are not covered by the regime, hence influencing trade and investment flows. Countries' terms of trade will shift, especially away from exporters of fossil fuels. And, if trading is permitted, sales of emission permits will redistribute income from one country to another.

Compliance

Inevitably, the question of international coordination brings up the issue of free riders or noncompliers. Since the signatories are sovereign states and there is no overarching disciplinarian, as there is (in principle) within countries, fines or economic sanctions are the possible means of forcing compliance.

It would be difficult to prohibit trade related to the emission of greenhouse gases without in effect prohibiting trade with the offending country, since carbon dioxide–producing energy is required for virtually all production. But prohibiting trade could impose costs not only on the offending countries but also on their trading

partners that might well exceed the likely costs of global climate change. The advantages of one international regime would be sacrificed for another. Even if the threat worked, in the sense that countries were induced to comply with the emission objectives and the threat therefore did not have to be exercised, its existence might induce some important countries—China comes to mind—to reduce their dependence on trade as a matter of policy to avoid the possible cost of sanctions in the future, and that too would represent a cost of compliance.

Chayes and Chayes (1995) have argued that treaty law need not rely predominantly on sanctions, either in theory or in fact, and that indeed in many cases sanctions are counterproductive. They conclude that most actual or apparent deviations from treaty provisions arise from ambiguity and indeterminacy of treaty language, from limitations on the capacity of governments to carry out their undertakings, or from major changes in circumstances from those prevailing at the time the treaty was ratified. Deliberate violation of treaties is rare, and when it occurs on important issues, major participants exert extreme pressure outside the treaty for resumption of compliance, as happened with the violation of the Non-Proliferation Treaty by Iraq and the threatened withdrawal from that pact by the Democratic Republic of Korea.

The key factors in ensuring compliance are a commitment to the treaty objectives plus a high degree of transparency in governmental actions. Many regulatory agreements have a potential free-rider problem; countries are more likely to adhere to the provisions if other governments are seen to be adhering to the provisions, so a regular system for monitoring and reporting on the activities and actions covered by the treaty is important.

These days, the very legitimacy of many governments arises from their responsibility for international relations and their integration into the community of nations. External and, increasingly, domestic pressure will usually keep governments from deliberately flouting internationally agreed behavior. The need to engage publics in the reduction of greenhouse gas emissions, however, raises the issue of whether governments have the capacity to carry out their international commitments. Taxes are easier to monitor than quantitative emission targets.

Incomplete Steps toward Mitigation

Understanding the processes of climate change and their social and economic impacts is a daunting task; the world's knowledge of ways to reduce emissions with minimal social disruption and to disseminate best practices remains highly imperfect. But knowledge can be advanced along a variety of fronts through the actions of individual countries, as well as through internationally sponsored research. Much of this activity is under way, especially research on improved energy efficiency and alternatives to fossil fuels, but governments should ensure that no promising idea is

languishing for lack of funding. Knowledge will not resolve some issues, howeve especially those concerning intergenerational distributional decisions and collecti aversion to risk. These inevitably must be resolved through public discussion an political negotiations.

Even in the absence of an effective international agreement, countries may sens bly take steps to reduce greenhouse gas emissions. Subsidies and tax advantages th encourage the consumption of fossil fuels, especially—but not only—in developir countries, can be removed. Countries can encourage more rapid diffusion of avai able best practice in energy use through schools and public awareness program Sometimes outdated regulations need to be changed; for example, allowable rates i public utility regulation are often based on investment in new generating capaci but not on investment in conservation of electricity. Governments can provide fundir for research in socially desirable new technologies and can, on their own, impo higher taxes on fossil fuels, devoting the revenues to reduction of other behavio distorting taxes. There are many reasons other than inhibiting global climate chang for adopting some or all of these measures: reduction of air pollution, reduction urban congestion, and enhancement of energy security (especially with respect imported oil). Reduction of greenhouse gas emissions would be a bonus, although conscious one.

Further, international lending institutions such as the World Bank and the re gional development banks are in the business of financing infrastructure projects i developing countries. Such infrastructure includes electric power generation and di tribution, transport systems, and other major power-using activities. Once infra structure is built, society adjusts to its long-lasting effects. Thus careful attentio should be paid now to the longer-run social and economic implications of the investments. The extent and character of waste emissions, including greenhouse gase should inform these investments, with special attention given to available best prac tice even when it is not considered by the principal contract-bidding firms. It important to seek viable alternatives to coal-fired electricity generating plants, in cluding a reexamination of the suitability of modern nuclear power technology wit respect to safety, cost, and waste disposal.

A sensitive issue arises when investments in low-emission or nonemitting tech nologies cost *more* than the least-cost technologies (taking into account initial inves ment, maintenance, and lifetime input and waste disposal requirements). Should th lending institutions nonetheless insist on the investment that is more friendly to th environment? If so, who should pay for the incremental cost? It seems reasonable t decline to finance infrastructure investments that are unnecessarily damaging to th environment and acceptable to ask the borrowing country to pay fully for the incre mental cost of any environmental benefits that accrue directly to it, but it is reason able to expect the international community to pay most or all of the incremental co

(depending on the income level of the borrowing country) associated with greenhouse gas emissions in cases in which the benefits will accrue to the world as a whole.

Contingency Planning

Many adverse developments *could* occur as a result of global climate change. It is much more difficult—if not impossible—to forecast with confidence what *will* actually happen. Some analysts have projected benign effects from global warming and easy adaptation to the adverse effects—especially for those whose income gives them room to maneuver. Thus developing countries understandably give higher priority to economic development than to averting climate change if the latter in any way inhibits development.

The great uncertainty about impacts, the prospect of serious gainers as well as losers, the high apparent cost of near-term actions to reduce emissions, and the need for eventual participation by countries with substantially different initial circumstances and hence greatly different priorities—all these factors make early action to stop the growth of greenhouse gas emissions, much less to lower them, highly problematic. Suppose the best guesses about climate change turn out to be too optimistic, or suppose that the forecasts are accurate but the international community is unable to reach agreement on costly, effective mitigation actions; or suppose that agreement is reached but countries prove unable to implement it. What will the community of nations do if experience reveals that climate change is severe and adverse? This possibility suggests the need for contingency planning to supplement research to develop cheap, low-emitting sources of energy and to satisfy human wants with lower requirements for energy. Such contingency planning can take two broad paths.

The first concerns how best to adapt to more serious climate change. It means not only pushing ahead with both the basic science and applied research for genetic engineering in many areas, especially agriculture, but also finding potential substitutes for possible useful species that may be lost That search could be supplemented by a systematic program for collecting, cataloguing, and storing genetic material, mainly but not exclusively from plants, in the form of seed banks and DNA.

The second concerns slowing global warming as rapidly as possible. One route involves sequestration and even withdrawal of greenhouse gases, mainly carbon dioxide, from the atmosphere on a scale at least equal to continuing emissions. That approach will involve good stack absorbers and storage depositories of carbon dioxide. But it also might involve mobilizing the biosphere. Rapidly growing trees could be planted on a massive scale (by using planes to drop seeds, for example), especially as climate change extends the areas that can support them. More unconventionally,

barren portions of the oceans could be fertilized with the requisite minerals (though mainly to be iron) so that microscopic carbon-loving plants can thrive.

A different approach would involve reducing the incidence of sunlight on Earth' surface, for example by placing reflecting surfaces in space or by increasing the al bedo by altering cloud formation or by placing particulates in the atmosphere (through jet engine exhaust or by using cannons or rockets). Other possibilities will no doub emerge over time. It is premature to commit to any particular method for rapi mitigation. Some suggestions will be impossibly expensive, and others will have un acceptable side effects, but it is important to encourage imaginative work on possibl emergency actions.

Notes

Richard N Cooper is professor of international economics at Harvard University

1. The principal tool of analysis for global climate change is a large computer model (Gener. Circulation Model, or GCM), of which several are in use, that attempts to model Earth's past an future climate as a function of received radiation, the characteristics of the atmosphere (such as cor centration of carbon dioxide), and the dynamics of climate formation

2. Excellent summaries of the scientific consensus and uncertainties about global climate chang can be found in IPCC (1996b) and in Wuebbles and Rosenberg (1998) More recent work sugges that increased ice in Antarctica induced by a warmer climate (due to higher precipitation ther would withdraw water from the oceans and thus reduce this projected rise in sea level, possibly by much as 50 percent (Thompson and Pollard 1997).

3 The European Union is treated as a single unit, with the maximum target reduction of 8 percer below emission levels of 1990 The United States agreed to a reduction of 7 percent and Japan to percent, and Russia agreed not to exceed its 1990 level

4 Fischer and Rosenzweig (1996) also find that global warming will increase global food produc tion by 2050, with carbon dioxide fertilization playing an especially important role

5 The comment by Kenneth Arrow misses my point Suppose, for the sake of argument, that a resources for additional investment are taken from consumption and none from investment, as Arro suggests The future is still best served if new investment is undertaken that has the highest rate return; an allocational mistake is made otherwise Hence, the appropriate discount rate for evaluatin potential projects is the rate of return on existing (marginal) investments

I am concerned with incremental decisions Arrow's approach would lead to a much higher rate *total* investment. Indeed, with enough potential projects yielding more than his preferred consum tion discount rate, it could lead logically to *all* output being used for investment—an absurd ou come Before it was reached, the consumption rate of discount would rise sharply, at least enough t keep consumption well above subsistence levels That would imply higher consumption rates of di count than Arrow either calculates or observes Another factor is relevant public investments must l covered, directly or indirectly, by taxation. Taxation in rich countries is already around 40 percent GDP (higher in Europe, lower in Japan and the United States). In pursuit of rational preferences, th voting public is not likely to support proposals that require much higher levels of public investmer with relatively low rates of return, which ultimately will require higher levels of taxation out of onl modestly higher income. This argument is weakened if the "investment" itself produces revenu which would be the case with a tax on carbon emissions.

References

Beckerman, Wilfred. 1996. *Through Green Colored Glasses* Washington, D C.: Cato Institute.

Chayes, Abram, and Antonia Handler Chayes. 1995 *The New Sovereignty Compliance with International Regulatory Agreements.* Cambridge, Mass Harvard University Press.

Cline, William R. 1998 "Comment on Lind and Schuler " In Nordhaus (1998. 97–103).

Council of Economic Advisers. 1996 *Economic Report of the President* Washington, D.C Government Printing Office.

———. 1998. "The Kyoto Protocol and the President's Policies to Address Climate Change· Administration Economic Analysis " Washington, D C Processed.

Darwin, Roy, Marinos Tsigas, Jan Lewandrowski, and A Raneses 1995 "World Agriculture and Climate Change· Economic Adaptations " Agricultural Economics Report AER-703 U S. Agriculture Department, Washington, D.C.

Dean, Andrew, and Peter Hoeller, eds. 1993. *The Costs of Cutting Carbon Emissions* Paris· Organisation for Economic Co-operation and Development.

Fischer, Cynthia, and Gunther Rosenzweig. 1996 "The Impacts of Climate Change, Carbon Dioxide, and Sulphur Deposition on Agricultural Supply and Trade An Integrated Impact Assessment." In N Nakicenovic, W. D Nordhaus, R Richels, and F L Toth, eds , *Climate Change· Integrating Science, Economics, and Policy* Laxenburg, Austria. International Institute for Applied Systems Analysis.

Intergovernmental Panel on Climate Change (IPCC) 1996a *Climate Change 1995 Economic and Social Dimensions of Climate Change Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge, U.K · Cambridge University Press

———. 1996b *Climate Change 1995 The Science of Climate Change Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge, U K. Cambridge University Press

Kram, Tom. 1998 "The Costs of Greenhouse Gas Abatement " In Nordhaus (1998)

Krause, Florentin, Wilfred Bach, and Jonathan Koomey 1992 *Energy Policy in the Greenhouse* New York: John Wiley

Mendelsohn, Robert, and James E Neumann, eds 1999 *The Impact of Climate Change on the United States Economy* New York Cambridge University Press

Meyer, William B , W Neil Adjer, Katrina Brown, Dean Graetz, Peter Gleich, John F Richards, and Antonio Maghalães 1998 "Land and Water Use " In Rayner and Malone, eds (1998, vol 2)

Moore, Thomas Gale 1998 *Climate of Fear Why We Shouldn't Worry about Global Warming* Washington, D.C.· Cato Institute

Nicholls, R. J , S P. Leatherman, K. C. Dennis, and C R Volanté 1995 "Impacts and Responses to Sea-Level Rise· Qualitative and Quantitative Assessments " In R. J Nicholls and S P Leatherman, eds , "The Potential Impact of Accelerated Sea-Level Rise on Developing Countries." *Journal of Coastal Research* 14(special issue)

Nordhaus, William D. 1994 *Managing the Global Commons The Economics of Climate Change* Cambridge, Mass. MIT Press

Nordhaus, William, D , ed 1998 *Economics and Policy Issues in Climate Change* Washington, D.C. Resources for the Future.

Pohl, Gerhard, and Dubravko Mihaljek. 1989 "Project Evaluation in Practice." World Bank. Processed

Psacharopoulos, George. 1985. "Returns to Education: A Further International Update and Implications." *Journal of Human Resources* 20(fall):583–604.

———. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22 1325–43.

Rayner, Steve, and Elizabeth L. Malone, eds. 1998. *Human Choice and Climate Change*. 4 vols Columbus, Ohio: Battelle Press.

Reilly, John. 1998. "Comment on 'Climate Change Damages.'" In Nordhaus (1998).

Schimmelpfennig, David, Jan Lewandrowski, John Reilly, Marinos Tsigas, and Ian Parry. 1996 "Agricultural Adaptation to Climate Change. Issues of Long-Run Sustainability." Agricultural Economics Report AER-740. U.S. Department of Agriculture, Economic Research Service, Washington, D.C.

Schmalensee, Richard. 1998. "Greenhouse Policy Architectures and Institutions." In Nordhaus (1998)

Stavins, Robert N. 1999 "The Costs of Carbon Sequestration: A Revealed-Preference Approach" *American Economic Review* 89(September):994–1009.

Thompson, Stanley L., and David Pollard. 1997. "Greenland and Antarctic Mass Balances for Present and Doubled Atmospheric CO_2 from the GENESIS Version-2 Global Climate Model." *Journal of Climate* 10(May):871–900.

Tol, Richard S. J. 1998 "Comment on 'Climate Change Damages.'" In Nordhaus (1998).

Wuebbles, Donald J., and Norman J, Rosenberg. 1998 "The Natural Science of Global Climate Change" In Rayner and Malone (1998, vol 2)

A Comment on Cooper

Kenneth J. Arrow

Richard Cooper has written an excellent survey on the economic implications of climate change, stressing the possibilities and limits of international policy. I comment here only on one part of the analysis, Cooper's choice of discount rates. I differ sharply with the basis of the analysis and wish to call attention to the extensive literature, dating back more than 40 years, on the choice of discount rates for public investment, which Cooper has disregarded.

Cooper refers to and then dismisses abruptly what may be called the *consumption* viewpoint. Investment is a sacrifice of consumption, and therefore the rate of return on a new investment should be at least equal to the implicit rate of return on consumption. (In this note, as in Cooper, all rates of return are real, not nominal.) This idea is hardly new; it is Marshall's "price of waiting." Bohm-Bawerk famously gave three grounds for the existence of a positive interest rate. First, if consumption is growing over time, the marginal utility of consumption must be falling; therefore, a sacrifice of consumption today must be compensated for by a greater increase in consumption in the future. Second, future consumption is automatically less valuable than the same consumption today, even if their marginal utilities are equal. Third, an increased lag of production behind inputs leads to an increase in production. The first two grounds together define the consumption rate of interest, as expressed in the formula (due, I believe, to Ragnar Frisch) $r = \rho + \theta g$, where ρ is the pure rate of time preference (corresponding to Bohm-Bawerk's second ground), θ is the elasticity of the marginal utility of consumption, and g is the rate of growth of consumption. The third is represented in modern language by the marginal productivity of capital, F_K, where $F(K)$ is output as a function of capital (taking labor and natural resources as given). The condition for optimal allocation over time, and the outcome of an intertemporal competitive equilibrium, is $r = F_K$.

Clearly, at an optimum, the rate of return on capital equals the consumption rate of interest, so that it would make no difference which rate is used for discounting. In fact, these two numbers seem to be very unequal. While estimating r is not straightforward, most estimates of the rate of return on consumption are on the order of 3 or 4 percent; the rate of return on capital is usually estimated (as by the U.S. Office of

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 173–75

Management and Budget) as 7–10 percent. Cooper also cites studies which argue that the rate of return on certain other public investments is 10–12 percent. He concludes that these rates represent alternative uses of capital and therefore that efficiency demands that the rates found in private investment or alternative public investments be used to evaluate climate change policies, in particular, abatement.

I believe this argument is fundamentally flawed, as pointed out some time ago by Eckstein (1957). There is no reason why the investment to prevent climate change must be drawn from other kinds of investment. It can be drawn from consumption and the consumption rate of interest is the alternative value of those resources. The more general situation is that the resources are drawn from both consumption and investment and therefore that the rate of return for evaluating public investment projects should be a weighted average of the consumption rate and the rate of return on capital. Since consumption is much larger than investment, it is reasonable to assume that the appropriate hurdle rate should be closer to the consumption rate.

The matter is further complicated by the fact that investments, public or private, spin off returns which, in turn, induce further saving and therefore investment, a point argued by Arrow and Kurz (1970). A full discussion of these points is well beyond the scope of this note, but the underlying structure of the approach has been set forth by Bradford (1975), and a complete explanation with a way of making the procedures operational is shown in Lind (1982).

Cooper quotes Maurice Scott as pointing out that the rate of return on risk-free securities is on the order of 4 percent. This observation is completely consistent with the previous estimates of the consumption rate of return. Cooper refuses to consider this rate as relevant, on the grounds that the investment should be made at the highest rate available. Of course, the correct conclusion is that *all* public investments that yield more than the consumption rate of return should be made, not that one should be made rather than the other. If the financing of public investments does not displace private investment, the rate of return on the latter is irrelevant.

The question that arises is, why does the market permit the discrepancy between the two rates of return? Two reasons are found in the literature. One is simply the corporate income tax, as stressed by Eckstein. The other is risk. For many reasons, including asymmetric information and moral hazard, the rate of return on private investments reflects a risk premium. The rate of return on public investments that we have been discussing is a riskless rate; the adjustment for uncertainty should come in the measurement of benefits and costs, not in the rate of return used in benefit-cost analysis. For these two reasons, the observed rate of return in the private sector is not the correct one for assessing public investment projects.

A final note: when considering investments for the very long future, the discussion has a strongly ethical component. I would argue that ethical preferences are subject to the same rules of rationality as private preferences, so that the *form* of the argument is much the same, although some of the parameters, particularly the rate of

pure time preference, may have different values ρ. But I believe the results will be quantitatively very similar.

Notes

Kenneth J. Arrow is Joan Kenney Professor of Economics, Emeritus and Professor of Operations Research, Emeritus at Stanford University

References

Arrow, Kenneth J , and Mordecai Kurz 1970 *Public Investment, the Rate of Return, and Optimal Fiscal Policy* Baltimore, Md . Johns Hopkins University Press for Resources for the Future

Bradford, David. 1975 "Constraints on Government Investment Opportunities and the Choice of Discount Rate " *American Economic Review* 65 887–95

Eckstein, Otto. 1957 "Investment Criteria for Economic Development and the Theory of Intertemporal Welfare Economics." *Quarterly Journal of Economics* 71 56–85

Lind, Robert C. 1982. "A Primer on the Major Issues Relating to the Discount Rate for Evaluating National Energy Options " In Robert C Lind, ed , *Discounting for Time and Risk in Energy Policy* Baltimore, Md.· Johns Hopkins University Press for Resources for the Future

Contagion: Understanding How It Spreads

Rudiger Dornbusch • Yung Chul Park • Stijn Claessens

Much of the current debate on reforming the international financial architecture is aimed at reducing the risks of contagion—best defined as a significant increase in cross-market linkages after a shock to an individual country (or group of countries) This definition highlights the importance of other links through which shocks are normally transmitted, including trade and finance. During times of crisis, the ways in which shocks are transmitted do seem to differ, and these differences appear to be important. Empirical work has helped to identify the types of links and other macroeconomic conditions that can make a country vulnerable to contagion during crisis periods, although less is known about the importance of microeconomic considerations and institutional factors in propagating shocks. Empirical research has helped to identify those countries that are at risk of contagion as well as some, albeit quite general, policy interventions that can reduce risks.

The financial turbulence that hit many East Asian countries in 1997 and then spread to other parts of the world continued unabated in 1998. Russia defaulted on its debt as confidence in global financial markets weakened. The turmoil roiled capital markets in industrial countries, dramatically altering the (relative) pricing of many financial instruments, and spilled over into speculative hedge-fund bets, leaving Long-Term Capital Management, a large U.S. hedge fund, facing near bankruptcy. The crisis subsequently hit Brazil, creating uncertainty about the country's ability to roll over its public sector debt, and spread to other emerging markets in Latin America and elsewhere.

International capital markets, particularly those in emerging markets, appear volatile, on both the downside and the upside. In the mid-1990s aggregate private capital flows into five crisis-affected East Asian countries (Indonesia, the Republic of Korea, Malaysia, the Philippines, and Thailand) averaged more than $40 billion annually, reaching a peak of about $70 billion in 1996. In the second half of 1997, more than $100 billion in short-term bank loans was recalled from these same five

The World Bank Research Observer, vol. 15, no 2 (August 2000), pp 177–97

countries, as currencies and stock markets there collapsed. Capital flows reversed themselves again in 1999, and stock markets rebounded sharply across the region a portfolio and other foreign investors channeled resources back, slowing the reform process in some countries. The turmoil triggered recessions in many developing countries, most notably in Latin America (Perry and Lederman 1998); altogether, two fifths of the global economy sank into recession in 1999, with the sharpest decline in gross domestic product concentrated in the developing world.

Neither the exact causes of this volatility nor the best international financial architecture for guiding the movement of international capital is yet known. Yet reducing volatility and contagion has been an important stated objective of recent reforms Fischer (1998), for example, notes two important reasons for revamping the international financial architecture and smoothing the global economy. First, the high degree of volatility of international capital flows to emerging markets and these markets' limited ability to deal with this volatility make the recipient country vulnerable to shocks and crises that are excessively large, frequent, and disruptive. Second, international capital markets appear to be highly susceptible to contagion. Thus proposals to reform the international financial architecture must be based on a thorough understanding of the causes and consequences of contagion.

Episodes of volatility in international capital markets had occurred before the Asian crisis; an example was the "tequila effect" that followed Mexico's December 1994 devaluation and mainly affected Latin American countries. At that time, the issue of financial contagion had not yet caught the attention of policymakers in either industrial or emerging-market countries (but see Kindleberger 1989). Since the East Asian crisis, however, policymakers and economists have engaged in considerable research to identify and analyze the causes of financial contagion.

Contagion is best defined as a significant increase in cross-market linkages after a shock to an individual country (or group of countries), as measured by the degree to which asset prices or financial flows move together across markets relative to their comovement in tranquil times An increase in comovement need not reflect irrational behavior on the part of investors. When one country is hit by a shock, liquidity constraints can force investors to withdraw funds from other countries. Because many financial transactions are conducted by agents rather than by principals, incentive issues also play a role in triggering volatility. A decision to pull funds from several countries can also reflect coordination problems among investors and insufficient mechanisms at the international level for dealing with countries' liquidity problems Distinguishing among these various forms of investor behavior is very difficult in practice.

Although it is hard to determine whether comovements are irrational or excessive empirical work has been able to document patterns in the vulnerability of countries to volatility and to identify possible channels through which contagion is transmit

ted. Trade links, regional patterns, and macroeconomic similarities make countries vulnerable to volatility. Volatility can be transmitted from a particular country to other countries through common creditors and through actions of investors operating in international financial centers. These regularities have helped to identify countries that are at risk of contagion. Less is known about the importance of microeconomic conditions and institutional factors (including the actions of specific financial agents) in propagating shocks.

Governments and the private sector, as well as international financial institutions, must take action to minimize and manage the risk of financial contagion. But the balance is unclear. Should individual countries bear the burden of improving their financial sectors and enhancing the transparency of data, or is there a need to reform the rules under which international investors operate? Does contagion always represent fundamental factors, or should countries simply have more access to liquidity support to withstand the pressures of contagion? For answers, we must first look at what is known about the causes and transmission of contagion.

Contagion and Its Causes

Contagion refers to the spread of market disturbances—mostly on the downside—from one country to the other, a process observed through comovements in exchange rates, stock prices, sovereign spreads, and capital flows In this article, we focus on contagion in emerging economies. The causes of contagion can be divided conceptually into two categories (Masson 1998; Wolf 1999; Forbes and Rigobon 2000; Pritsker 2000). The first category emphasizes spillovers that result from the normal interdependence among market economies. This interdependence means that shocks, whether of a global or local nature, can be transmitted across countries because of their real and financial linkages. Calvo and Reinhart (1996) term this type of crisis propagation "fundamentals-based contagion." These forms of comovements would not normally constitute contagion, but if they occur during a period of crisis and their effect is adverse, they may be expressed as contagion. Most empirical work seeks to explain the degree of comovements and the mechanisms for transmitting them—for example, how and under what conditions a speculative attack on a single currency is spread to other currencies on the basis of various fundamental relationships.

The second category involves a financial crisis that is not linked to observed changes in macroeconomic or other fundamentals but is solely the result of the behavior of investors or other financial agents. Under this definition, contagion arises when a comovement occurs, even when there are no global shocks and interdependence and fundamentals are not factors. A crisis in one country may, for example, lead investors to withdraw their investments from many markets without taking account of differences in economic fundamentals. This type of contagion is

often said to be caused by "irrational" phenomena, such as financial panics, herd behavior, loss of confidence, and increased risk aversion. But because these phenomena can be individually rational and still lead to a crisis, it is helpful to discuss each category in detail.

Fundamental Causes

Fundamental causes of contagion include macroeconomic shocks that have repercussions on an international scale and local shocks transmitted through trade links, competitive devaluations, and financial links.

COMMON SHOCKS. Studies identify various global shocks that can trigger market adjustments in an international context. A common global cause, such as major economic shifts in industrial countries and changing commodity prices, can trigger crises in—or large capital inflows to—emerging markets. Changes in U.S. interest rates have been identified with movements in capital flows to Latin America (Calvo and Reinhart 1996; Chuhan, Claessens, and Mamingi 1998). The strengthening of the U.S. dollar against the yen in 1995–96 was an important factor in the export downturn in East Asia and the subsequent financial difficulties there (Corsetti, Pesenti, and Roubini 1998; Radelet and Sachs 1998a, 1998b). In general, a common shock can lead to comovement in asset prices or capital flows.

TRADE LINKS AND COMPETITIVE DEVALUATIONS. Local shocks, such as a crisis in one economy, can affect the economic fundamentals of other countries through trade links and currency devaluations. Any major trading partner of a country in which a financial crisis has induced a sharp currency depreciation could experience declining asset prices and large capital outflows or could become the target of a speculative attack as investors anticipate a decline in exports to the crisis country and hence a deterioration in the trade account.

Competitive devaluations can be another channel for transmitting contagion. Devaluation in a country hit by a crisis reduces the export competitiveness of the countries with which it competes in third markets, putting pressure on the currencies of other countries, especially when those currencies do not float freely. According to Corsetti and others (1999), a game of competitive devaluation can induce a sharper currency depreciation than that required by any initial deterioration in fundamentals. In addition, the noncooperative nature of the game can result in still greater depreciation compared with what could have been attained in a cooperative equilibrium. If market participants expect that a currency crisis will lead to a game of competitive devaluation, they will naturally sell their holdings of securities of other countries, curtail their lending, or refuse to roll over short-term loans to borrowers in those countries. This theory gains some credence from the fact that during the East

Asian crisis in 1997, exchange rates depreciated substantially even in economies such as Singapore and Taiwan, China, which did not necessarily appear vulnerable to a speculative attack on the basis of their fundamentals.[1]

FINANCIAL LINKS. Economic integration of an individual country into the world market typically involves both trade and financial links. Thus a financial crisis in one country can lead to direct financial effects, including reductions in trade credits, foreign direct investment, and other capital flows abroad. For example, firms in East Asia that are linked to, say, Thailand by trade, investment, and financial transactions would be adversely affected if a crisis were to limit the ability of Thai firms to invest abroad, extend credit, and so on. Thus a financial crisis in Thailand would rationally be reflected in other countries, leading, for example, to comovements in asset prices and capital flows.

Investors' Behavior

The spread of a crisis depends on the degree of financial market integration. If a country is closely integrated into global financial markets, or if the financial markets in a region are tightly integrated, asset prices and other economic variables will move in tandem. The higher the degree of integration, the more extensive could be the contagious effects of a common shock or a real shock to another country. Conversely, countries that are not financially integrated, because of capital controls or lack of access to international financing, are by definition immune to contagion. In this sense, financial markets facilitate the transmission of real or common shocks but do not cause them. The actions of investors that are ex ante individually rational as well as collectively rational, even though they lead to volatility and may require policy changes, should be grouped under fundamental causes.

It can be argued, however, that investors' behavior, whether rational or irrational, allows shocks to spill over from one country to the next. The literature differs on the scope of rational versus irrational investor behavior, both individually and collectively. It is useful to start with a classification of types of investor behavior (see also Pritsker 2000). First, investors can take actions that are ex ante individually rational but that lead to excessive comovements—excessive in the sense that they cannot be explained by real fundamentals.[2] Through this channel, which can broadly be called investors' practices, contagion is transmitted by the actions of investors outside the country, each of whom is behaving rationally. Conceptually, this type of investor behavior can be further sorted into problems of liquidity and incentives and problems of informational asymmetry and market coordination. Second, cases of multiple equilibrium, similar to those in models of commercial bank runs, can imply contagious behavior among investors. Third, changes in the international financial system, or in the rules of the game, can induce investors to alter their behavior after an initial crisis.

LIQUIDITY AND INCENTIVE PROBLEMS. One form of rational behavior by individuals relates to liquidity and other constraints on lenders or investors. For example, the sharp currency depreciation and the decline in equity prices in Thailand and other economies affected early in the East Asian crisis resulted in large capital losses for some international institutional investors. These losses may have induced investors to sell off securities in other emerging markets to raise cash in anticipation of a higher frequency of redemptions. Liquidity problems may also face commercial banks whose lending is concentrated in particular regions. Suppose there is a single common creditor country with a heavy regional exposure, such as Japan in East Asia or the United States in Latin America. If banks from the common creditor country experience a marked deterioration in the quality of their loans to one country, they may attempt to reduce the overall risk of their loan portfolios by reducing their exposure to other high-risk investments elsewhere, possibly including other emerging markets in the region.

The incentive structure for individual financial agents can also create a tendency to sell off several markets at the same time. For example, an initial crisis may induce investors to sell off their holdings in other emerging countries because of their tendency to maintain certain proportions of a country's or a region's stock in their portfolios. As a result, equity and other asset markets in a range of emerging economies would also lose value, and the currencies of these economies would depreciate significantly. Schinasi and Smith (2000), for example, demonstrate that the value-at-risk models used by many commercial banks explain why financial institutions and other investors may find it optimal to sell most high-risk assets when a shock affects one of those assets. Although this type of behavior is individually rational, it can lead to overall adverse outcomes.[3] Garber (1998) analyzes the possible unpleasant dynamics associated with the use of unregulated financial derivatives in weak institutional settings.

Countries whose financial assets are widely traded in global markets and whose domestic financial markets are more liquid may be more vulnerable to financial contagion (Kodres and Pritsker 1998; Calvo and Mendoza forthcoming). Further, because global diversification of financial portfolios involves the cross-market hedging of macroeconomic risks, countries in which asset returns exhibit a high degree of comovement with a crisis-affected country in tranquil periods will be more vulnerable to contagion (Kaminsky and Reinhart 1998b).

These liquidity constraints and incentive structures could be important for all types of investors dealing with emerging markets. But it is possible that particular institutional investors—open-end emerging-market mutual funds, hedge funds, and proprietary traders—are especially susceptible to this type of behavior. Leveraged investors, such as hedge funds and banks facing margin calls, are more likely to confront liquidity problems in the wake of a crisis and be forced to sell their asset holdings in other markets. Managers of open-end funds may also need to raise li

quidity in anticipation of future redemptions by investors. Faced with these problems, both leveraged investors and open-end-fund managers are likely to keep those assets whose prices have already collapsed and whose secondary markets have become less liquid and sell other assets in the portfolio. By doing so, investors cause other asset prices to fall, and the original disturbance can spread across different financial instruments and markets. The financial turmoil in the fall of 1998, when spreads on U.S. corporations rose from a normal level of 100 basis points to almost 200 basis points, suggests that these types of spillovers need not be limited to emerging markets but can also affect a broad spectrum of markets and borrowers.

INFORMATION ASYMMETRIES AND COORDINATION PROBLEMS. Another cause of contagion relates to imperfect information and differences in investor expectations. In the absence of better information to the contrary, investors may believe that a financial crisis in one country could lead to similar crises in other countries. A crisis in one country may then induce an attack on the currencies of other countries in which conditions are similar. This type of behavior can reflect rational as well as irrational behavior. If a crisis reflects and reveals weak fundamentals, investors may rationally conclude that similarly situated countries are also likely to face such problems; such reasoning helps explain how crises become contagious. This channel presumes, of course, that investors are imperfectly informed about each country's true characteristics and thus make decisions on the basis of some known indicators, including those revealed in other countries, which may or may not reflect the true state of the subject country's vulnerabilities The information investors use may include the actions of other investors, which brings us to the effects of informational asymmetries on investor behavior.

Investors often do not have a full picture of the condition of every country as it affects their return on investment. In part, this limitation reflects the cost of gathering and processing information. Calvo and Mendoza (forthcoming) show that in the presence of information asymmetries, the fixed costs involved in gathering and processing country-specific information could lead to herd behavior, even when investors are rational. In their model, financial investors can be divided into two groups: informed and uninformed. Given the fixed cost of gathering and processing information, most small investors simply cannot afford to collect and process country-specific information individually (see also Agenor and Aizenman 1998). Instead, uninformed investors may find it less costly and therefore advantageous to follow the investment patterns of informed investors. In making asset choices, uninformed investors may then take into account portfolio decisions made by better-informed investors because such decisions provide useful market information

Both informed and uninformed investors may tend to seek new information from those investors who acted earlier to adjust their portfolios. Thus if informed investors move to pull out of a country, the information cascade will lead less-informed

LIQUIDITY AND INCEN ... mation and follow the informed investors, thereby
relates to liquidity ... (Scharfstein and Stein 1990; Wermers 1995
sharp currency de ... he tendency to herd may increase as the num-
economies affec ... can be placed grows and the range of investors
some internat ... athering and processing country-specific in-
to sell off sec ... that an increase over time in herd behavior
frequency ... khchandani, Hirshleifer, and Welch 1992,
lending ... f optimal portfolio diversification that be-
count ... row (Calvo and Mendoza forthcoming)
Sta ... rding over time is that as investors have
m ... a reputation has become relatively more
... expensive to follow the herd. Because some inves
...nagers, may be more concerned about maintaining a repu-
...on the performance of their portfolios, relative to that of a given
portfolio, than about their absolute performance, the risks of cascading be-
havior may be particularly high among institutional investors (see Kim and We 1997 for foreign exchange trading). Thus an individual institutional investor may refrain from acting first, even if market developments favor a new portfolio, for fear of losing his or her reputation if the decision should prove to be wrong. To be on the safe side, individual investors may follow the herd. All these outcomes involve behavior that is individually rational (albeit constrained) but that nevertheless can cause financial volatility.

ced with these prob- likely to keep those markets have be- investors cause ross different 1998, when to almost emerg

MULTIPLE EQUILIBRIUMS. A more general explanation of contagion based on investors' behavior involves changes in expectations that are self-fulfilling in financial markets subject to multiple equilibriums. In this framework, contagion occurs when a crisis in one emerging market causes another emerging-market economy to move or jump to a bad equilibrium, characterized by a devaluation, a drop in asset prices capital outflows, or debt default. In Diamond and Dybvig's (1983) model of bank runs, it is rational for individual depositors to either hold funds in the bank or withdraw funds, depending on the actions of all other depositors. The equilibrium result can be a bad outcome, that is, a run on the bank, or a good outcome, in which depositors keep their money in the bank. In an economic crisis, the result analogous to a bank run would be a sudden withdrawal of funds from a country sparked by investors' fears that unless they act quickly they will be too late to claim the limited pool of foreign exchange reserves.

Some observers argue that contagion is a consequence of sudden shifts in market expectations and confidence. Formal analytical models of multiple equilibriums have been developed to explain recent experience in emerging markets (Gerlach and Smets 1995; Jeanne 1997; Masson 1998). Such models, of course, do not lend themselves easily to empirical tests because the move or jump can be triggered by many factors

some of which may appear to be fundamental causes. Drazen (1999), for example, shows that political factors may have played a role in the contagion during the 1992–93 Exchange Rate Mechanism crisis. And, of course, such changes in equilibrium are not limited to emerging markets but can also play a role in volatility and contagion in domestic financial markets.

CHANGES IN THE RULES OF THE GAME. Finally, contagion may result if investors change their assessment of the rules under which international financial transactions occur. The Russian default in 1998, for example, increased concern that other countries might follow similar unilateral policies regarding the treatment of foreign private creditors or that international financial institutions might not bail such creditors out as expected. The discussion on the international financial architecture itself following the East Asian financial crisis may have caused changes in the way investors viewed the rules of the game and weighed the odds of official bailouts. This concern is often alleged to have caused the turbulence in 1998 in Brazil (see Calvo 1998; Park 1998; Dornbusch 1999). Other reasons could include concern about the supply of funds from international lenders of last resort In late 1998, for example, the International Monetary Fund (IMF) found itself called on to rescue so many countries that economists wondered whether it would be able to deal with many more liquidity crises. Thus a liquidity crisis in one country could trigger a run on other countries out of fear that the last eligible country would be out of luck.

Empirical Evidence of Contagion

Empirical examination of the evidence on contagion has focused mainly on comovements in asset prices rather than on "excessive" comovements in capital flows or disturbances in real markets. We discuss tests under the following categories: correlation of asset prices; conditional probabilities of a currency crisis; changes in volatility; comovements of capital flows and rates of return; and other tests.

Correlation of Asset Prices

The asset price tests measure the correlation among different economies in interest rates, stock prices, and sovereign spreads (Forbes and Rigobon 1999 survey the recent literature). A marked increase in correlations is considered evidence of contagion. Most of these studies find evidence of large comovements in a variety of asset returns, although there is less agreement on whether such comovements increase in the wake of a crisis. Several studies suggest that the Mexican crisis in 1994 was contagious. Calvo and Reinhart (1996) find that the comovement of weekly returns on equities and Brady bonds in emerging markets in Asia and Latin America was higher

after the Mexican crisis than before. Frankel and Schmukler (1998) show evidence that the prices of country funds in Latin America and East Asia displayed greater comovement with Mexican country funds after the crisis than before. Valdés (1997) confirms that the movements of secondary-market debt prices and credit ratings show that the Mexican crisis was contagious in Latin America. Agenor, Aizenman, and Hoffmaister (1999) report that the Mexican crisis had a sizable effect on movements in domestic interest rate spreads (and output) in Argentina. Baig and Goldfajn (1998) show that the cross-country correlations among currencies and sovereign spreads in Indonesia, Korea, Malaysia, the Philippines, and Thailand increased significantly during the East Asian crisis (from July 1997 to May 1998) compared with other periods.

A marked increase in correlations among markets in different countries may, however, not be sufficient proof of contagion. If markets are historically cross-correlated, a sharp change in one market will naturally lead to changes in other markets, and correlations during crises could increase appreciably. Forbes and Rigobon (2000) show that as volatility increases following a crisis, an increase in correlation could simply be a continuation of strong transmission mechanisms that exist in more stable periods. They also show that an increase in correlations of asset prices may result when changes in economic fundamentals, risk perception, and preferences are correlated, without any additional contagion. Because of this endogeneity, estimation of correlations must control both for comovement in these variables during normal times and for the effects of fundamentals in order to be able to identify pure contagion.

In practice, it is impossible to adjust for the effects of increases in volatility and endogeneity (as well as omitted variables) without making some more restrictive assumptions. Some papers have done so. Forbes and Rigobon (1999) investigate the evidence of contagion during the 1987 U.S. stock market crash, the 1994 Mexican peso crisis, and the 1997 East Asian crisis using daily data for stock indexes of up to 28 industrial countries and emerging markets. They show that correlation coefficients across multicountry returns are not significantly higher during crises, if one properly corrects for the problems of endogenous variables, omitted variables, and changes in the variance of residuals. Arias, Hausmann, and Rigobon (1998) also find only limited evidence of contagion.

In a test on the Exchange Rate Mechanism crisis, however, Favero and Giavazzi (2000) estimate a structural model of the behavior of European interest rates and find evidence of contagion in interest rate residuals even after controlling for normal interdependence. Using an autoregressive model, and thus controlling to some degree for structural relationships, Park and Song (2000) show that the Southeast Asian crisis did not directly trigger the crisis in Korea but that its fallout to Taiwan played an important role in the Korean crisis (see also Connolly and Wang 1998 and Tan 1998 for comovements of stock prices in Asia; Doukas 1989 for sovereign spreads).

Conditional Probabilities

Another way to control for the role of fundamentals is to study conditional correlation or probabilities, rather than raw correlations, and thus use a narrower definition of contagion. The most commonly used methodology, introduced by Eichengreen, Rose, and Wyplosz (1996) and Sachs, Tornell, and Velasco (1996), examines whether the likelihood of crisis is higher in a given country when there is a crisis in one or several other countries. This literature builds on studies in single-country crisis prediction (see Dornbusch, Goldfjan, and Valdés 1995; Sachs, Tornell, and Velasco 1996). Berg and Pattillo (1999) review this literature, and Goldstein, Kaminsky, and Reinhart (2000) provide a more general exposition of early warning systems.

The research involves estimating the probability of a crisis conditional on information on the occurrence of crisis elsewhere, taking into account fundamentals or similarities. One advantage of this definition of contagion is that it readily allows for statistical tests of its existence. These tests can also try to investigate the channels through which contagion may occur, distinguishing, among others, trade and financial links. Eichengreen, Rose, and Wyplosz (1996), using a probit model and a panel of quarterly macroeconomic and political data covering 20 industrial economies from 1959 through 1993, show that the probability of a domestic currency crisis increases with a speculative attack on a currency elsewhere and that contagion is more likely to spread through trade linkages than through macroeconomic similarities. Using a similar methodology, De Gregorio and Valdés (2000) conduct an extensive test of spillovers of the 1982 debt crisis, the 1994 Mexican crisis, and the Asian crisis using indexes of exchange rate pressures over three- and twelve-month horizons, real exchange rate movements, and changes in credit ratings.[4] They find that the Mexican crisis was the least contagious, while the Asian crisis was as contagious as the 1982 crisis (note that their methodology does not allow them to determine whether spillovers represent normal comovements or contagion). Importantly, they find that both debt composition and exchange rate flexibility limit the extent of contagion, whereas capital controls do not appear to curb it

Taking an even longer perspective, Bordo and Murshid (2000) examine the record of financial crises over the past 120 years and the evidence of contagion in several macroeconomic variables. They find that the core countries of the prewar and interwar gold standards (the United Kingdom and the United States) appear to be important in disseminating shocks to the rest of the world but that such patterns actually appear to be weaker during crises. In contrast, after 1973, Bordo and Murshid find that countries that are otherwise not correlated show considerable comovement in asset prices during crises. They also find, however, that the volatility in correlation coefficients can be quite high; they are therefore reluctant to interpret the increase in correlations during recent periods as evidence of contagion, especially in light of the

Forbes and Rigobon (2000) finding that such increases might be normal. On the whole, these tests find no solid evidence that contagion has been increasing over time.

Glick and Rose (1998) apply a similar approach to five episodes of currency crises and 161 countries and find that trade linkages are important in propagating a crisis. They argue that contagion tends to be regional rather than global because trade tends to be more intraregional than interregional (see also Diwan and Hoekman 1999). Kaminsky and Reinhart (1998a) find that in terms of conditional probabilities, information on a large share of crisis countries in the sample increases the ability to predict a crisis elsewhere, particularly on a regional level. Their study further supports evidence that contagion has been primarily a regional phenomenon (see also Calvo and Reinhart 1996; Frankel and Schmukler 1998; Kaminsky and Schmukler 1999).

The evidence on the trade channel as an explanation of the regional nature of contagion appears more relevant to Latin America than to East Asia. Kaminsky and Reinhart (1998a) find a high probability that a crisis will spread through third-party linkages among Latin American countries (Brazil, Colombia, Mexico, and Venezuela), while similar linkages are not significant in East Asia. Brazil, Colombia, Mexico and Venezuela have the largest share of bilateral trade with the United States among Latin American countries. Baig and Goldfajn (1998) analyze the trade matrix of East Asian countries and find that trade linkages among those countries are weak. They argue that trade linkages were not important in the expansion of the crisis in East Asia in 1997. Alba and others (1999) investigate the effects of competitive devaluations and argue that these alone could not have explained the large depreciation of other regional currencies after the Thai devaluation.[5] In transition economies, Gelos and Sahay (2000) find that correlations in pressures on the exchange market can be explained by direct trade linkages but not by measures of other fundamentals. They also find that market reactions following the Russian crisis look very similar to those observed in other regions during turbulent times. Tests thus find strong evidence that contagion is related to trade links and has a regional character.

Kaminsky and Reinhart (1998b) find that the probability of contagion increases when the crisis is associated with the common creditor channel. Indonesia, Malaysia, and Thailand are heavily dependent on Japanese commercial bank lending; a crisis in one or two of these countries spread to all three. Similar results are found in Latin America, where the conditional probability of a crisis in one country when several others are in crisis is estimated to be as high as 78 percent. Latin American countries obtain a large portion of credit from U.S. commercial banks. Analogous effects appear for other types of investors. Using data on closed-end country funds, Frankel and Schmukler (1998) test whether adverse shocks from the Mexican crisis were transmitted directly, or indirectly through financial markets based in New York. They find that Wall Street spread

the Mexican crisis to East Asian countries but did not play a role in its transmission to other Latin American countries.

Volatility Spillover

Another approach estimates spillovers in volatility—that is, cross-market movements in asset prices. Edwards (1998) examines Mexico's interest rate increase in 1994 and finds strong evidence of contagion from Mexico to Argentina but not from Mexico to Chile. Park and Song (1999) test volatility spillover among foreign exchange markets during the crisis period and find that the effects of the crises in Indonesia and Thailand were transmitted to the Korean foreign exchange market but that the Korean crisis did not reinfect the two Southeast Asian countries. These studies did not control for fundamentals and thus did not distinguish between a pure contagion and one based on fundamentals.

Capital Flows

Capital flows can offer the best insight into the transmission of contagion, but few tests of their comovements have been conducted. Van Rijckeghem and Weder (2000) test the role of bank lending and the effect of a common lender by examining capital flows to 30 emerging markets. In the Mexican and Russian crises, they find that the degree to which countries obtained funds from common bank lenders was a fairly robust predictor of both disaggregated bank flows and the incidence of a currency crisis. Froot, O'Connell, and Seasholes (2000) study the behavior of portfolio flows into and out of 44 countries from 1994 through 1998 They find strong evidence that price increases encourage portfolio flows and that price declines lead to reduced flows. They also find that regional factors such as common creditors appear to be increasingly important over time, suggesting that the actions of institutional investors could be a channel for transmission of shocks.

In an analysis of portfolios of mutual funds, Kaminsky, Lyons, and Schmukler (forthcoming) find that emerging-markets funds exhibit positive momentum. That is, they systematically buy winners and sell losers in both crisis and noncrisis periods, with one difference: contemporaneous momentum (buying current winners and selling current losers) is stronger during crises, whereas lagged momentum (buying past winners and selling past losers) is stronger during noncrisis periods. Contemporaneous momentum was at its strongest point during the 1994 crisis in Mexico. Importantly, Kaminsky, Lyons, and Schmukler find that mutual fund managers use contagion strategies; that is, they sell assets from any country when crisis hits another—strong evidence that contagion is transferred through the actions of portfolio investors. Choe, Kho, and Stulz (1999) find that foreign portfolio investors did not add to volatility (see also Kim and Wei 1999; Stulz 1999).

Other Tests

Most empirical papers find that macroeconomic weaknesses can provoke contagion because they make a country vulnerable to a crisis. Similarities in macroeconomic weaknesses can also lead to crisis because these signals are considered sorting devices and thus may induce a shift in investors' expectations. Ahlumawia (2000) attempts to separate the two effects and finds that after controlling for the direct effect of weaknesses, macroeconomic similarities can play a proximate role in contagious currency crises by coordinating investor shifts. A study of the behavior of the local lending activities of domestic- and foreign-owned banks in Argentina and Mexico reveals that foreign-owned banks may have had a stabilizing influence on overall credit growth in the banking sector, potentially reducing both countries' vulnerability to crisis (Goldberg, Dages, and Kinney 2000). There have been few tests using structural models to explain the degree of spillovers in real and financial markets. One is the application of a full trade model for crisis-affected East Asian economies (Abeysinghe 2000). Although transmission through trade played an important role, Abeysinghe found that the immediate economic contractions were largely a result of direct shocks attributable to pure contagion.

Implications and Reform Options

The empirical findings show that fundamentals help predict spillovers and that trade links are important factors as well. Common creditor and other links through financial centers transmit volatility from one country to another at a particular point This work thus helps to identify those countries that are at risk of a spillover of volatility. Less is known about the importance of microeconomic conditions and institutional factors—including actions of specific financial agents and the various channels that induce spillovers—in propagating shocks. As a result, it has been difficult to attribute the spillovers to contagion. Importantly, the degree of spillover does not appear to have increased over time, and there are many similarities in the empirical regularities across periods and countries.

These findings suggest that comovements are unavoidable and that fundamental factors are important. To reduce the risks of financial contagion, reforms will thus be necessary. Many of these are of a general nature, such as reductions in fiscal and current account deficits, better management of exchange rates, improvements in financial sector services, enhancement of data transparency, and the like. Many economists have proposed, and some have analyzed, specific policy reform options to deal with contagion. Stiglitz and Bhattacharya (2000) argue, for example, that disclosure requirements may not be needed because markets can and do provide optimal incentives for disclosure. They also argue that under certain circumstances, information

disclosure could exacerbate fluctuations in financial markets and precipitate a financial crisis. Bushee and Noe (1999), looking at U.S. equity markets, find that improved disclosure by firms increases the volatility of their stock prices because the seemingly reduced information asymmetry and increased liquidity of the market attract more transient investors. Here Furman and Stiglitz (1998) point to the fact that even countries such as Sweden, with good regulation and supervision and transparent financial markets, have had financial crises.

Many economists also agree that although improved standards (for data disclosure, regulation, supervision, and corporate governance) could have prevented the buildup of vulnerabilities and reduced the risk of currency crises, they are only a first step. Improved implementation and surveillance are necessary as well. For example, Hawkins and Turner (2000), who analyze the role of prudential and other standards for financial institutions, stress implementation issues and predict that many developing countries will continue to have difficulty complying with what are essentially industrial-country standards.

For these reasons, several observers have argued for the use of prudential controls, particularly for financial institutions, to limit the risk of sudden capital outflows. Many countries already limit the maturity mismatches on foreign exchange liabilities and assets, monitor internal risk management systems of financial institutions, and issue sanctions for poor systems. Tightening could mean putting limits on the net open positions that financial institutions can take in foreign currency markets, as well as imposing limits on the amount of gross foreign currency liabilities (as a fraction of total liabilities or as a ratio to equity). Guidelines on internal risk management systems can be issued, and financial institutions can be more intensely monitored in this area. A further precautionary measure would require banks to hold more liquid foreign exchange assets relative to total foreign exchange liabilities than they are required to hold on domestic currency liabilities. And, finally, capital controls on (some type of) inflows at the country level might be useful to prevent the buildup of vulnerabilities; there is much less agreement in this area, however.

Specific reforms to the rules under which international investors operate are less apparent. There have been calls for limits on the operations of hedge funds, and revisions to the way in which commercial banks have to hold assets against short-term loans to emerging markets. But so far, no proposals specifically aimed at curbing the role of investors in contagion have emerged, let alone been agreed on. More discussion has occurred on the need to enhance liquidity support to withstand pressures of contagion, perhaps though an international lender of last resort or standstills on payments following a crisis. Clearly, whatever reforms are implemented, liquidity crises will still arise; thus a good part of the debate on the international financial architecture has focused on improving ways for dealing with the crises. In an analysis of the supply of international liquidity, Chang and Majnoni (2000) stress that liquidity provisions entail a tradeoff: liquidity provisions conditioned on certain poli-

cies and applied at penalty rates can deepen the possibility of a full crisis. At the same time, moral hazard concerns call for conditions and higher rates. Some new facilities—the Supplemental Reserves Facility, the Contingent Credit Lines of the IMF, the guarantee facility of the World Bank, and private sector facilities—are set up ex ante, which may reduce these concerns. They may also induce foreign investors to avoid generating a level of debt that may place the economy in a fragile situation

Conclusion

Economists still do not know precisely what factors make countries vulnerable to contagion or the exact mechanisms through which it is transmitted at any given time. Although empirical evidence suggests that commercial banks and mutual funds can play a role, separating rational from irrational investor behavior is difficult in theory and in practice, as is determining whether irrational investor behavior is the sole source of contagion. Individually rational but collectively irrational behavior and (perceived) changes in the international financial system are likely to continue to have an influence. Further research—whether theoretical or empirical—on the role of international financial agents and the international financial system may shed light on these aspects. Such research could help identify characteristics that make countries vulnerable to contagion and could contribute to the development of specific policy prescriptions to reduce the risks of contagion, manage its impact, and help economies recover as efficiently as possible. In the meantime, it will be difficult to determine whether any measures—beyond strengthening the international financial architecture—can reduce the risks of contagion specifically.

Notes

Rudiger Dornbusch is professor of economics at the Massachusetts Institute of Technology, Yung Chul Park is professor of economics at Korea University, and Stijn Claessens is lead economist in the Financial Sector Policy Group of the World Bank An earlier version of this paper was prepared for discussion at the World Institute for Development Economics Research workshop on financial contagion held at the World Bank, June 3–4, 1999, and reflects comments from participants.

1 An interesting question is whether Singapore and Taiwan, China, let their currencies depreciate to maintain export competitiveness or to conserve foreign reserves. Corsetti and others (1999) argue that these two economies were able to defend the original parities with their massive holdings of reserves and thus to withstand irrational withdrawal but were concerned about a loss of competitiveness It can also be argued, however, that their decision to float their currencies was motivated by their efforts to fend off possible speculative attacks driven both by arbitrary shifts in expectations and by the reaction of panicky and irrational investors Although the response may have been rational and optimal in either case, in that the perceived welfare costs of maintaining a stable exchange rate might have been too high, the contagion aspects and policy implications underlying the two rationales are quite different.

2. Investors can follow strategies that are ex ante irrational given their own preferences and the behavior of other investors. Although one cannot rule out the likelihood that this category is large, its lack of conceptual definition makes it difficult to analyze.

3 In a related argument, Goldfajn and Valdés (1997) find that when foreign investors withdraw deposits and loans, asset prices decline and asset markets become illiquid. Banks and other financial institutions thus risk failure because they cannot readily liquidate their assets The liquidation problem may cause a run on these intermediaries themselves, provoking a banking or confidence crisis, and could lead to a speculative attack on the currency as foreign investors withdraw and convert their investments into foreign exchange Such crises can spread to other countries when international investors are forced to sell off their positions in other national markets to make up for the liquidity shortage caused by the crisis in one country.

4 Caramazza, Ricci, and Salgado (1999) investigate the East Asian, Mexican, and Russian crises using an approach similar to that of Eichengreen, Rose, and Wyplosz (1996). They find that these crises do not differ much. Fundamentals, including trade spillovers, common creditors, and financial fragility, are highly significant in explaining crises, while exchange rate regimes and capital controls do not seem to matter

5 In contrast, Baig and Goldfajn (1998) find large trade links among East Asian countries, which could explain some spillover based on reduced demand for intraregional exports (see also Huh and Kasa 1997)

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems

"The Contagion Conference" refers to the conference "International Financial Contagion: How It Spreads and How It Can Be Stopped," sponsored by the World Bank, the International Monetary Fund, and the Asian Development Bank, February 3–4, 2000, Washington, D.C. Papers from the conference are available at <http //www worldbank org/research/interest/confs/past/papersfeb3-4/papers htm>.

Abeysinghe, Tilak 2000 "Thai Meltdown and Transmission of Recession within ASEAN4 and NIE4 " The Contagion Conference

Agenor, Pierre-Richard, and Joshua Aizenman 1998. "Contagion and Volatility with Imperfect Credit Markets." *IMF Staff Papers* 45(June) 207–35

Agenor, Pierre-Richard, Joshua Aizenman, and A Hoffmaister 1999 "Contagion, Bank Lending Spreads, and Output Fluctuations." World Bank Institute, Washington, D C. Processed.

Ahlumawia, Pavan 2000 "Discriminating Contagion An Alternative Explanation of Contagious Currency Crises in Emerging Markets " IMF Working Paper WP/00/14 International Monetary Fund, Washington, D.C

Alba, Pedro, Amar Bhattacharya, Stijn Claessens, Swati R Ghosh, and Leonardo Hernandez 1999 "Volatility and Contagion in a Financially Integrated World Lessons from East Asia's Recent Experience " In Gordon de Brouwer with Wisarn Pupphavesa, eds , *Asia Pacific Financial Deregulation* London and New York. Routledge.

Arias, Eduardo Fernandez, Ricardo Hausmann, and Roberto Rigobon. 1998. "Contagion on Bond Markets: Preliminary Notes." Massachusetts Institute of Technology, Sloan School of Management, Cambridge, Mass Processed

Baig, T., and I Goldfajn 1998. "Financial Market Contagion in the Asian Crisis " IMF Working Paper WP/98/155. International Monetary Fund, Washington, D C

Banerjee, Abhijit. 1992. "A Simple Model of Herd Behaviour." *Quarterly Journal of Economics* 107 797–817.

Berg, Andrew, and Catherine Pattillo 1999. "Are Currency Crises Predictable? A Test." *IMF Staff Papers* 46 (June):107–38.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Changes as Information Cascades." *Journal of Political Economy* 1:992–1020

Bordo, Michael D., and Antu P. Murshid. 2000. "The International Transmission of Financial Crises before World War II· Was There Contagion?" The Contagion Conference.

Bushee, Brian, and Christopher Noe 1999. "Unintended Consequences of Attracting Institutional Investors with Improved Disclosure" Working Paper 00-33. Harvard Business School, Cambridge, Mass. Processed.

Calvo, Guillermo 1998. "Understanding the Russian Virus (with Special Reference to Latin America)" University of Maryland, Department of Economics, College Park, Md. Processed.

Calvo, Guillermo, and Enrique Mendoza Forthcoming "Rational Herd Behavior and the Globalization of the Securities Market." *Journal of International Economics.*

Calvo, Sara, and Carmen Reinhart. 1996 "Capital Flows to Latin America. Is There Evidence of Contagion Effect?" In Guillermo Calvo, Morris Goldstein, and Eduard Hochreiter, eds, *Private Capital Flows to Emerging Markets after the Mexican Crisis* Washington, D.C Institute for International Economics.

Caramazza, Francesco, Luca Ricci, and Ranil Salgado 1999. "Trade and Financial Contagion in Currency Crises" IMF Working Paper WP/00/55. International Monetary Fund, Washington, D.C Processed.

Chang, Roberto, and Giovanni Majnoni. 2000. "International Contagion. Implications for Policy" The Contagion Conference

Choe, Hyuk, Bong-Chan Kho, and René Stulz 1999 "Do Foreign Investors Destabilize Stock Markets? The Korean Experiences in 1997" *Journal of Financial Economics* 54(2):227–64

Chuhan, Punam, Stijn Claessens, and Nlandu Mamingi 1998 "Equity and Bond Flows to Asia and Latin America The Role of Global and Country Factors" *Journal of Development Economics* 55·439–63

Connolly, Robert, and Albert Wang 1998 "International Equity Market Comovements. Economic Fundamentals or Contagion?" University of North Carolina, Kenan-Flagler Business School, Chapel Hill, N C Processed

Corsetti, Giancarlo, Paolo Pesenti, and Nouriel Roubini 1998 "What Caused the Asian Currency and Financial Crises? Part I A Macroeconomic Overview" NBER Working Paper 6833. National Bureau of Economic Research, Cambridge, Mass

Corsetti, Giancarlo, Paolo Pesenti, Nouriel Roubini, and Cédric Till 1999 "Competitive Devaluations. A Welfare-Based Approach." NBER Working Paper 6889 National Bureau of Economic Research, Cambridge, Mass

De Gregorio, José, and Rodrigo O Valdés 2000 "Crisis Transmission. Evidence from the Debt Tequila, and Asian Flu Crises" The Contagion Conference

Diamond, Douglas, and Philip Dybvig 1983 "Bank Runs, Deposit Insurance, and Liquidity" *Journal of Political Economy* 91·401–19

Diwan, Ishac, and Bernard Hoekman 1999 "Competition, Complementarity, and Contagion in Asia" In Pierre-Richard Agenor, Marcus Miller, David Vines, and Axel Weber, eds., *The Asian Financial Crisis Cause, Contagion, and Consequences* Cambridge, U.K.. Cambridge University Press.

Dornbusch, Rudiger. 1999 "Brazil's Incomplete Stabilization and Reform." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass Processed.

Dornbusch, Rudiger, Ilan Goldfajn, and Rodrigo Valdés. 1995. "Currency Crises and Collapses." *Brookings Papers on Economic Activity* 2:219–95

Doukas, John. 1989 "Contagion Effects on Sovereign Interest Rate Spread" *Economic Letters* 29:237–41.

Drazen, Allan. 1999. "Political Contagion in Currency Crises" NBER Working Paper 7211 National Bureau of Economic Research, Cambridge, Mass

Edwards, Sebastian 1998. "Interest Rate Volatility, Capital Controls, and Contagion." NBER Working Paper 6756. National Bureau of Economic Research, Cambridge, Mass.

Eichengreen, Barry, Andrew Rose, and Charles Wyplosz 1996 "Contagious Currency Crises." NBER Working Paper 5681 National Bureau of Economic Research, Cambridge, Mass

Favero, Carlo A, and Francesco Giavazzi 2000 "Looking for Contagion Evidence from the 1992 ERM Crisis." The Contagion Conference.

Fischer, Stanley 1998 "Reforming the International Monetary System" Draft paper prepared for the David Finch Lecture, Melbourne, Australia, presented November 9 Available at <http://www imf.org/external/np/speeches/1998/110998 htm>

Forbes, Kristin, and Roberto Rigobon 1999 "No Contagion, Only Interdependence. Measuring Stock Market Comovements" NBER Working Paper 7267 National Bureau of Economic Research, Cambridge, Mass.

——— 2000. "Measuring Contagion: Conceptual and Empirical Issues." The Contagion Conference

Frankel, Jeffrey, and Sergio Schmukler 1998 "Crisis, Contagion, and Country Fund Effects on East Asia and Latin America." In Reuven Glick, ed , *Managing Capital Flows and Exchange Rates Perspectives from the Pacific Basin* New York· Cambridge University Press

Froot, Kenneth A , Paul G. J O'Connell, and Mark S. Seasholes 2000. "The Portfolio Flows of International Investors, I" The Contagion Conference

Furman, Jason, and Joseph Stiglitz 1998. "Economic Crises Evidence and Insights from East Asia" *Brookings Papers on Economic Activity* 2 1–135

Garber, Peter. 1998. "Derivatives in Capital Flows" NBER Working Paper 6623 National Bureau of Economic Research, Cambridge, Mass

Gelos, R Gaston, and Ratna Sahay 2000 "Financial Market Spillovers in Transition Economies." The Contagion Conference

Gerlach, Stefan, and Frank Smets 1995 "Contagious Speculative Attack." *European Journal of Political Economy* 11.5–63

Glick, Reuven, and Andrew Rose 1998 "Contagion and Trade Why Are Currency Crises Regional?" NBER Working Paper 6806 National Bureau of Economic Research, Cambridge, Mass

Goldberg, Linda, Gerard Dages, and Daniel Kinney 2000 "Lending in Emerging Markets Foreign and Domestic Banks Compared." The Contagion Conference

Goldfajn, Ilan, and Rodrigo O. Valdés 1997 "Capital Flows and the Twin Crises· The Role of Liquidity" IMF Working Paper WP/97/87 International Monetary Fund, Washington, D C

Goldstein, Morris, Graciela Kaminsky, and Carmen Reinhart 2000 *Assessing Financial Vulnerability An Early Warning System for Emerging Markets* Washington, D C Institute for International Economics.

Hawkins, John, and Philip Turner 2000 "International Financial Reform Some Issues" The Contagion Conference

Huh, Chan, and Kenneth Kasa 1997 "A Dynamic Model of Export Competition, Policy Coordination, and Simultaneous Currency Collapse" Pacific Basin Working Paper Series PB97-08. Federal Reserve Bank of San Francisco, San Francisco, Calif

Jeanne, Olivier. 1997. "Are Currency Crises Self-Fulfilling?" *Journal of International Economics* 43 263–86.

Kaminsky, Graciela, and Carmen Reinhart. 1998a. "Financial Crisis in Asia and Latin America Then and Now" *American Economic Review* 88(May, *Papers and Proceedings*):444–48.

———. 1998b. "On Crises, Contagion, and Confusion." George Washington University, Economics Department, Washington, D.C Processed.

Kaminsky, Graciela, and Sergio Schmukler. 1999. "On Booms and Crashes: Stock Market Cycles and Financial Liberalization." George Washington University, Economics Department, Washington, D C. Processed.

Kaminsky, Graciela, Richard Lyons, and Sergio Schmukler. Forthcoming. "Economic Fragility, Liquidity, and Risk. The Behavior of Mutual Funds during Crises." *Journal of International Economics*

Kim, Jungshik, and Shang-Jin Wei 1997. "The Big Players in the Foreign Exchange Market Do They Trade on Information or Noise?" NBER Working Paper 6256 National Bureau of Economic Research, Cambridge, Mass

Kim, Woochan, and Shang-Jin Wei. 1999 "Foreign Portfolio Investors before and during a Crisis" NBER Working Paper 6968. National Bureau of Economic Research, Cambridge, Mass.

Kindleberger, Charles. 1989 *Manias, Panics, and Crashes* Rev. ed New York: Basic Books

Kodres, Laura, and Matthew Pritsker. 1998 "A Rational Expectations Model of Financial Contagion" Finance and Economics Discussion Paper 9848 Board of Governors of the Federal Reserve System, Washington, D C

Masson, Paul 1998 "Contagion. Monsoonal Effects, Spillovers, and Jumps between Multiple Equilibria" IMF Working Paper WP/98/142 International Monetary Fund, Washington, D.C

Park, Yung Chul 1998 "Financial Crisis in Korea and Its Lessons for Reform of the International Financial System" In Jan Joost Teunissen, ed., *Regulatory and Supervisory Challenges in a New Era of Global Finance* The Hague: FONDAD

Park, Yung Chul, and Chi-Young Song 1999 "East Asian Financial Crisis. One Year After." *IDS Bulletin* 30(1) 93–103 Institute of Development Studies, University of Sussex, Brighton, U K

——— 2000 "Financial Contagion in the East Asian Crisis With Special Reference to the Republic of Korea." The Contagion Conference.

Perry, Guillermo, and Daniel Lederman 1998 "Financial Vulnerability, Spillover Effects, and Contagion" World Bank, Chief Economist's Office, Latin America, Washington, D.C Processed

Pritsker, Matt 2000 "The Channels for Financial Contagion" The Contagion Conference

Radelet, Steven, and Jeffrey Sachs. 1998a "The East Asian Financial Crisis Diagnosis, Remedies Prospects." *Brookings Papers on Economic Activity* 1 1–90

——— 1998b "The Onset of the East Asian Currency Crisis" NBER Working Paper 6680 National Bureau of Economic Research, Cambridge, Mass.

Sachs, Jeffrey, Aaron Tornell, and Andres Velasco 1996. "Financial Crises in Emerging Markets The Lessons from 1995." *Brookings Papers on Economic Activity* 1 147–215

Scharfstein, David, and Jeremy Stein 1990. "Herd Behavior and Investment." *American Economic Review* 80.465–79

Schinasi, Garry J., and R. Todd Smith. 2000. "Portfolio Diversification, Leverage, and Financial Contagion" The Contagion Conference

Shiller, Robert 1995. "Conversation, Information, and Herd Behavior." *American Economic Review* 85 (May, *Papers and Proceedings*).181–85.

Stiglitz, Joseph E., and Amar Bhattacharya Forthcoming "Underpinnings for a Stable and Equitable Global Financial System. From Old Debates to a New Paradigm " *Annual World Bank Conference on Development Economics 1999.* Washington, D C World Bank

Stulz, René M. 1999. "International Portfolio Flows and Security Returns " Working Paper 99 3 Ohio State University, Charles A. Dice Center for Research in Financial Economics, Columbus, Ohio. Available at <http://www.cob.ohio-state edu/fin/faculty/stulz>.

Tan, Jose A. 1998 "Contagion Effects during the Asian Financial Crisis Some Evidence from Stock Price Data." Pacific Basin Working Paper Series PB98-06 Federal Reserve Bank of San Francisco, San Francisco, Calif

Valdés, Rodrigo. 1997. "Emerging Markets Contagion Evidence and Theory " Massachusetts Institute of Technology, Economics Department, Cambridge, Mass Processed

Van Rijckeghem, Caroline, and Beatrice Weder. 2000 "Financial Contagion Spillovers through Banking Centers." The Contagion Conference

Wermers, Russ 1995. "Mutual Fund Herding and the Impact on Stock Prices " *Journal of Finance* 54(April):581–622

Wolf, Holger 1999. "International Asset Price and Capital Flow Comovements during Crisis The Role of Contagion, Demonstration Effects, and Fundamentals " Paper presented at the World Bank/IMF/WTO conference on "Capital Flows, Financial Crises, and Policies," April 15–16, Washington, D C

Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries

Deon Filmer • Jeffrey S. Hammer • Lant H. Pritchett

Recent empirical and theoretical literature sheds light on the disappointing experience with implementation of primary health care programs in developing countries. This article focuses on the evidence showing two weak links in the chain between government spending for services to improve health and actual improvements in health status. First, institutional capacity is a vital ingredient in providing effective services When this capacity is inadequate, health spending, even on the right services, may lead to little actual provision of services. Second, the net effect of government health services depends on the severity of market failures—the more severe the market failures, the greater the potential for government services to have an impact. Evidence suggests that market failures are the least severe for relatively inexpensive curative services, which often absorb the bulk of primary health care budgets. A companion paper, available from the authors (see p. 219), offers a perspective on how government funds can best be used to improve health and well-being in developing countries. It gives an alternative view of appropriate public health policy, one that focuses on mitigating the characteristic market failures of the sector and tailoring public health activities to the government's ability to deliver various services.

The concept called primary health care, or health for all, enshrined at an international conference at Alma Ata (now Almaty, Kazakhstan) in 1978, has dominated much of the discussion of health policy in developing countries for the past two decades. The broad health problems this concept encompasses are pressing, and the solutions it calls for seem obvious. Yet in many developing countries the public budget for health is principally absorbed by public hospitals staffed by doctors expensively trained at public expense who use costly medical technologies to treat conditions of the urban elite, while in those same countries children die from diseases that could have been treated for a few cents or avoided altogether with basic hygienic practices. In developing countries in 1995, more than 9 million children under five

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 199–224

years old—more than the entire population of Sweden or Zambia—died deaths that could have been avoided.[1]

Individual examples, however, show that primary health care efforts can be successful even in very poor regions. Kerala, a state in India with annual per capita income of only $1,254 in purchasing power parity dollars, has an infant mortality rate of only 31 per 1,000 live births (Agrawal and others 1996; International Institute for Population Studies 1995). This rate is not only 40 percent lower than that in Punjab, another Indian state with twice Kerala's income, but 35 percent lower than that in Brazil, which has more than four times Kerala's per capita income (Macro International 1997; World Bank 1999). Infant mortality in Shanghai, China, is lower than in Manhattan, and the recorded infant mortality rate of 16 in Jamaica is lower than that of African-Americans in the United States. Ceara, one of the poorest states in Brazil, reduced infant mortality by 36 percent in just a few years through an aggressive government program (Tendler and Freedheim 1994).

This combination of experiences has led to a strong consensus among public health specialists who focus on developing countries. They argue that the existing allocation of health expenditures toward curative care in secondary and tertiary facilities, such as hospitals and clinics to which patients are referred, is inappropriate and that a reorientation of government efforts toward primary health care would bring both health gains and cost savings. In this consensus primary health care is typically defined by what it is *not:* it is neither secondary nor tertiary curative care but could be *all* other activities related to health, from nutrition, to sanitation, information, and education, to clinic-based curative care. Even more ambitious definitions include empowerment and social equality (Decosas 1990).

Although the images and statistics that motivate primary health care appear compelling, the gains have rarely been demonstrated in practice. Individual initiatives, such as the expansion of immunization campaigns or the recent efforts to combat river blindness, have been effective in a single region or for a single disease, but the data do not typically show that other aspects of public health care have much impact on health. The implementation of primary health care programs has created a new set of images: empty rural health clinics without staff, drugs, or working equipment, poor people bypassing free primary public clinics to pay for services from private providers; government-supplied drugs for sale on the black market. This disappointing experience raises the question: what was missing from the seemingly compelling logic? There appear to be several weak links in the chain that prevent public primary health care from being delivered to more people. We believe that economists bring to bear two perspectives on this question that are useful for understanding health policy: choices and incentives.

First, the expected impact of primary health care was too often calculated as if health status were entirely a technical affair and individuals were the passive recipients of government action. But individuals are guided by their own knowledge and

resources in judging the quality of their health care (and that of their children). Incorporating their choices into the analysis can change both the expected overall impact of these programs and the importance of various actions, because service delivery and overall effectiveness both depend on the demand for specific services, the price of the services, and the existing (and potential) supply in the private sector.

Second, primary health care advocates assumed that the public sector could deliver whatever the government (or some international forum) decided *ought* to be delivered. In practice, the quality of public health services has ranged from excellent to truly horrific. While an ideal, well-run network of community workers and rural health clinics might have a dramatic effect on health status, the real policy issue is what the public sector is actually capable of providing. Often, health service failures result from a systemic mismatch between the traditional civil service incentive structure and the tasks required in the health sector.

A Simple Framework for Analyzing Public Spending on Health

In general, public spending influences health status by lowering the effective price of health-enhancing inputs (whether information on food cleanliness or a heart transplant). How it does so depends on four distinct mechanisms:

- *Composition of public spending.* The effect of an increase in total public spending on health depends on how that increase is allocated across health inputs. An equal increase in spending on all inputs will have a very different effect from an increase in spending on only the most effective interventions.
- *Output of the public sector.* The next issue is the magnitude of effective health services created by the public spending. Whether a government decides to build a clinic, spray for malaria, mobilize community outreach workers, or buy X-ray machines, it can be more or less effective at translating that expenditure into a real supply of health services. This efficacy will have a country-specific component common to all activities, as well as an activity-specific component, since government might be better adapted to performing some activities than others.
- *Net impact of public sector supply on overall consumption.* Even when government funds are available to provide the service in question, one must assess whether the change in the effective price faced by consumers will translate into increased overall consumption. Public spending on a particular service may not be cost-effective in improving health if additional consumption in the public sector crowds out—in whole or in part—equally effective services from nongovernmental providers. The size of this crowding-out effect depends on how individual demand and private sector supply respond to changes in the price, travel time, convenience, or quality of services available.

- *The health production function.* Different health inputs are more or less effective in improving health in ways determined by biological and medical facts. This health production function—that is, determinations about which treatments are effective in eliminating which cancers, which vaccines are potent over what period, how micronutrients affect susceptibility to diseases—is the domain of health care professionals. Economists typically prefer to remain agnostic about the particulars of the production function, and in many cases sensible recommendations about public policy need not inquire into the production function

Within this framework, the consensus argument is that increases in public spending on primary health care are effective in improving aggregate measures of health status, while curative services at secondary and tertiary levels are not.[2] This view can rationalize an increase in funds for primary health care as well as a reallocation of the health budget toward primary health care activities. The argument, however, typically jumps directly from the assumption that if the government spends on the right things, patients will receive the right things. Primary health care advocates rely on evidence showing that when evaluated as interventions, primary health care elements of the health production function are typically more cost-effective than secondary or tertiary activities. But the actual impact of public spending is the product of all four terms described above: allocation of the budget, efficacy of the public sector, market impact on consumer demand for services, and the actual effect of the health services on health. If any one of these is low, the total impact will be low, and the middle two are the weak links.[3]

Has Primary Health Care Been Effective at the National and Local Levels?

Has public spending on health and, more particularly, on primary health care promoted good health? If so, there should be empirical regularities at both the national and local levels. First, given a level of total health expenditures, more spending on primary health care activities and greater access to primary health care services should be associated with lower aggregate mortality. Second, at the local level (household, village) greater access to primary health care facilities should reduce mortality. Third, projects that develop primary care facilities should reduce mortality. None of these regularities finds much support in the data.

It might seem odd that we do not review the literature evaluating primary health care per se. As one recent review of that literature highlights, however, the literature does not permit such an evaluation. Fox (1995:2) analyzed the 87 articles published since 1980 in four major health journals (*Health Policy, Health Policy and Planning,*

International Journal of Health Services, and *Social Science and Medicine*) that included the words "primary health care" in the Medline subject category:

> While there are questions about parts of PHC [primary health care] (e.g. cultural acceptability of community participation ideals, selection processes for PHC workers, community financing critiques, the capture of program planning or implementation by funding agencies, or the ability to have health for all in socially unjust societies), there was no serious questioning in the literature of PHC as a desirable way for Ministries of Health to spend their money. . . . Most of the program evaluations or topic evaluations don't show great technical expertise in evaluation methodology being applied to PHC, but a fair number of articles discuss how such an evaluation would be done, or why it would be difficult.

A similar problem affects the empirical evaluation of "health status." The World Health Organization (WHO) defines health as "a state of complete physical, mental, and social well-being, and not merely the absence of disease or infirmity" (WHO 1988). Although this is an attractive definition, it is subjective and hard to assess. Mortality is easy to measure, whereas morbidity is not, but (fortunately for people, but unfortunately for research) mortality is rare (except among the very young and very old). Moreover, because death occurs rarely, and only once for each person, it is difficult to study at the individual level. So even though interest focuses on health over an individual's lifetime, empirical studies tend to view infant (or child) mortality or life expectancy as proxies for health status and to rely on mortality or life expectancy rates at an aggregate level (district, province, and country) for comparisons.[4]

Comparisons at the National Level

Cross-national studies of health status have come to a fair consensus on two points. First, socioeconomic characteristics explain nearly all of the variation in mortality rates across countries. A recent econometric study of mortality among children under age five shows that virtually all the cross-national variation in child mortality can be explained by six variables: average gross domestic product (GDP) per capita, a measure of the distribution of income, the level of female education, a dummy variable for predominantly Muslim countries, an index of ethnolinguistic diversity, and a set of five dummy variables for regions (Filmer and Pritchett 1999).

Second, total public spending on health has had much less impact on average health status than one might have expected and certainly less than one could have hoped for. Although the lack of data on public spending has, until recently, limited direct examination of the issue, Musgrove (1996:44) summarizes studies of the effect of public spending on health as follows: "Multivariate estimates of the determi-

nants of child mortality give much the same answer [as his results on life expectancy]: income is always significant, but the health share in GDP, the public share in health spending, and the share of public spending on health in GDP never are." A recent paper on global health from the Population Reference Bureau downplays the role of aggregate spending, stating that "these resources, however, usually do not adequately measure the *effectiveness* or availability of services" (Ratzan, Filerman, and LeSar 2000:36, emphasis added).

Other recent studies support these findings. Using instrumental variables to account for data and endogeneity problems, Filmer and Pritchett (1999) find that public expenditure on health as a share of GDP is a small and statistically insignificant determinant of child mortality. Estimates find that doubling public spending from 3 to 6 percent of GDP would improve mortality by only 9–13 percent. Bidani and Ravallion (1997) show that public spending has a large impact on the health status of the poor, but they estimate the effect of public spending on aggregate health status (of the poor and nonpoor taken together) to be quite small.

We say the consensus is "fair" because there are some exceptions to these two points. First, Anand and Ravallion (1993) could not reject the restriction that average income affected health status only insofar as it affected the level of poverty, highlighting the fact that increased income among the poor is the most effective variable in improving health. The very high correlation between average income and poverty still implies, however, that a country's average income explains most of the variation in health status. Second, an influential study (Preston 1980), based on data from 1940 to 1970, emphasized the low explanatory power of socioeconomic variables. Studies using more recent data, however, including one by Preston (1986), who relied on data from the 1970s, are unanimous about the high explanatory power of socioeconomic basics such as average income and female education.

The cross-national evidence has always been absent or ambivalent on whether health status is improved by a greater commitment to or greater spending on primary health care (or both). First, despite the power of socioeconomic conditions to determine health status, there are still outliers, such as Kerala in India, Sri Lanka, and Costa Rica, whose achievements are potentially replicable. Although these examples suggest possibilities, it was never very clear whether the success stemmed directly from characteristics of their health systems or was instead a social or political phenomenon. Kunstadter (1985:234), participating at a seminal conference that affirmed support for primary health care (at least of the selective variety) on the basis of case studies of the outliers, commented:

> The four case studies [China, Costa Rica, Kerala State, and Sri Lanka] involve societies in which low mortality has been reached without high per capita income. Situations in which low income continues to be associated with high mortality or high income is associated with high mortality were not

considered, nor have we searched systematically for other societies in which relevant social characteristics of the four successful cases are repeated, to see what happened to mortality. Thus the policy prescriptions are relatively weak.

Moreover, although China's barefoot doctors are famous, it is not obvious that their success can be attributed to a primary health care strategy—nor even that it was related to medical care. The largest declines in infant and child mortality occurred in the 1950s and the first half of the 1960s, before the introduction of the barefoot doctors in late 1965. Further, when the program was abandoned (ironically, at about the time of the Alma Ata conference), health status did not deteriorate (Liu and others 1995). In 1986 Sri Lanka spent 70 percent of its public monies on hospitals, substantially higher than the 56 percent average for comparable countries in South Asia (Griffin 1992). Our regressions for this article suggest that when either the fraction of the budget spent on "local health services" or the "fraction of population with access to local health services" is added to a regression of mortality on income and other country socioeconomic characteristics, it has no statistically or empirically significant impact (Filmer, Pritchett, and Hammer 1998).[5] Although this is weak evidence *against* the effectiveness of primary health care, we know of no published cross-national evidence that lends *support* to primary health care.

Outcomes at the Local Level

A second empirical regularity that would support primary health care would be evidence that the availability of primary-level health facilities or community health workers had a demonstrable impact on the local health status of individuals and communities. Unfortunately, the empirical results about the effect on health status of proximity to hospitals, doctors, and, in particular, public sector clinics, health centers, and rural health workers are, at best, mixed

A major problem in assessing this evidence is that governments may have systematically located clinics in underserved areas where health status is worst. As a result, the comparison of health status in localities with and without clinics would understate the true impact of such clinics. Conversely, if the government places clinics in villages where the demand is greatest, these sites may be located in areas where health would have been good in any case, and comparisons of sites with and without a facility would overstate the impact of clinics.

Recent studies assess the effect of access to services on child or infant mortality, using methods that take into account the bias resulting from the placement of facilities Frankenberg (1995) controls for placement effects by comparing randomly matched births from two different age cohorts in a village in Indonesia. She finds that the presence of a maternity clinic or a doctor reduced mortality but that the

presence of a health worker other than a doctor *increased* the probability of death (although this finding was statistically significant only at the 10 percent level). Pitt, Rosenzweig, and Gibbons (1993) compare matched districts in Indonesia and find that villages in a district served by a health center had higher mortality rates, whereas those with a family planning clinic had lower mortality rates; in both cases, the findings are statistically insignificant. When she does not control for selective placement, Frankenberg (1995) finds that the presence of more maternity clinics and health workers (insignificantly) reduced mortality, while more doctors (insignificantly) increased mortality. In uncorrected estimates, Pitt, Rosenzweig, and Gibbons (1993) find that the presence of both health centers and family planning clinics raised mortality rates (insignificantly in the case of health centers).

There are many studies that do not control for selective placement. After controlling for the potential endogeneity of facility usage, Panis and Lillard (1994) find that delivering a baby within a health care institution in Malaysia (the likelihood of which increases with the availability of such facilities) reduced the probability that the baby would subsequently die; puzzlingly, they also find that prenatal care (insignificantly) increased the probability that the child would subsequently die. Benefo and Schultz (1996) find that proximity to a clinic reduced child mortality in Côte d'Ivoire but appeared to increase mortality (statistically insignificantly) in Ghana. Lavy and others (1996) find that proximity to health facilities significantly increased mortality in rural Ghana (although the availability of child services in the closest clinic was significantly positively related to survival). In Malaysia DaVanzo (1984) finds that distance to medical care was (insignificantly) related to infant mortality (conditional on birthweight) but that low birthweight was correlated with the distance to care. Hammer, Nabi, and Cercone (1995) find that public medical facilities in Malaysia were unrelated to mortality, whereas Hossain (1989) reports that the presence of a dispensary and a family planning clinic lowered mortality in Bangladesh.

Rosenzweig and Schultz (1982) find that rural health posts, municipal public and private clinics, dispensaries, and mobile care units were not significantly related to child mortality in rural Colombia. In urban areas, hospitals, clinics, and family planning centers tended to reduce mortality, but this result was not consistent across groups sorted by the age of the mother. Rosenzweig and Wolpin (1982) use data from rural India to find that villages in a district with a family planning clinic and those with a dispensary were associated with lower child mortality but that villages with any "other health facility (health centers, nursing homes, etc. . . .)" were associated with higher mortality. The large National Family Health Survey of India shows no relation between health centers or subcenters and child mortality (World Bank 1998a).

Sastry (1995) finds that the number of general health facilities is associated with higher mortality in northeastern Brazil but with lower mortality in southern and southeastern Brazil (both effects are statistically insignificant). In a study of census

districts in Brazil, Paes de Barros, Da Costa, and Mendoça (1998) find that the availability of public health personnel had no effect on mortality.

Overall, the econometric evidence that clinics have *any* effect on health is mixed. Even where such evidence exists, the impact is too small to explain much of the variation in health status.

Evaluating the Effect of Health Care Programs on Child Mortality

In 1977 a very intensive maternal and child health and family planning program was introduced in a set of treatment villages in the Matlab region of Bangladesh (with a nearby set of villages serving as comparators). In the treatment area mothers and children were visited every 15 days by a female health worker who provided guidance on family planning. Detailed records were kept in both areas. Although mortality among children fell (Muhuri and Preston 1991), analysts exploiting the controlled design of this project attribute the decline almost exclusively to measles immunization (Koenig, Fauveau, and Wojtyniak 1991; Menken and Phillips 1990).

A separate project that assessed the impact of health care services for children in Narangwal, India, found that infant mortality dropped 6 percentage points over the three years of the project (from 96 per 1,000) in the treatment area compared with a 1 percentage point increase in the control area (from 107 per 1,000), but the difference was insignificantly different from zero (Taylor and Singh n.d.). The change in the project villages resulted from a 10 percentage point increase in perinatal mortality, a 4 percentage point increase in neonatal mortality, and a 16 percentage point decrease in postneonatal mortality. In the control villages, perinatal mortality rose only 1 percentage point, neonatal mortality fell 21 percentage points, and postneonatal mortality increased 49 percentage points. The last figure leads the authors to question the resulting 1 percentage point increase reported for overall infant mortality. The authors' explanation is that this huge fluctuation was "due to incorrect age classification, . . . since in the control villages . . . investigators who had no access to exact birth information needed to rely on the age at death given by the mother of the family" (p. iv.D.7).

These empirical results suggest that enhancing health outcomes is not simply a matter of providing additional funds or increasing access to primary health care services and facilities. The framework discussed earlier provides two likely explanations for the negligible impact of public spending. Each of these explanations leads to different policy implications. First, the impact of primary health care provision depends on the effectiveness of the service provided. If this is the problem, the solution is an improvement in the quality of the service. The question then becomes: what policy levers are available to increase quality? Second, the impact of the service depends on individual choice and the market for health, that is, private demand may vary by disease condition and the response by private suppliers to public interven-

tion. The question in this case is, what are the factors that affect private supply ; individual demand?

Public Sector Spending and the Creation of Effective Health Services

Primary health care may have little impact on health status not because such car unimportant but because, in practice, the efficacy of government health interv tions may be low. Readers in industrial countries without personal experience r find it difficult to appreciate just how poor the quality of public sector services be. Anecdotal reports describe situations marked by corruption and mismanagem In one low-income country, for example, a prominent newspaper accused the he ministry of misappropriating $50 million of donor financing. The next day the m istry accused the newspaper of exaggeration for failing to make it clear that this : million was misappropriated over a period of three years, not in a single year as newspaper report implied.

In a second case, a client survey of women who had delivered a child in the p two years at rural health centers in the Mutasa district of Zimbabwe (Mtemeli 19 listed the most frequently cited disadvantages of giving birth in an institution. Th were "ridiculed by nurses" for not having baby clothes (22 percent), "maternity fe (16 percent), "nurses ordered mothers to wash used linen soon after delivery" percent), and "nurses hit mothers during delivery" (!) (13 percent). Interestin when the nurses themselves were asked why they thought mothers did not delive health institutions, they cited, first, the distance to the facility and transport pr lems (20 percent); second, inability to pay (14 percent); and, third, *harassmen nursing staff or fear of nurses* (11 percent).

Moreover, in nearly every country one can find rural health clinics that have drugs, although government- (or donor-) financed medicines are readily available the black market. For example, more than 70 percent of the government supply drugs disappeared in Guinea in 1984 (Foster 1990). Various studies in Camero Tanzania, and Uganda estimated that about 30 percent of publicly supplied dr was misappropriated; in one case as much as 30 to 40 percent of the public sup was "withdrawn for private use" by staff (World Bank 1994a).

Evidence on Quality

Alderman and Lavy (1996) recently reviewed several empirical studies that examine the link between the quality of and the demand for public facilities; among these studies were Akin, Guilkey, and Denton (1995); Lavy and Germain (1994); Lavy and others (1996); Mwabu, Ainsworth, and Nyamete (1993); and Thomas, Lavy,

and Strauss (1996). Even though the measures of quality are not always satisfactory, the findings confirm that demand is responsive to service characteristics. An example of the problematic nature of some of the measures is that the absence of various types of drugs is often used to indicate inadequacies in public health delivery. Shortages, however, could be caused by high levels of demand, and inferences about the causal relationship are difficult to draw (a problem generally acknowledged by the authors cited above). In addition, there may be important discrepancies between de facto and de jure measures of quality, a result highlighted by Thomas, Lavy, and Strauss (1996), who show the different effects on health outcomes of the actual number of staff members compared with the official number.

Both the ineffectiveness of low-level public sector health clinics and the element of individual choice are highlighted in the phenomenon called "bypassing," which occurs when people choose to bypass the closest public facility in favor of either more costly private facilities or higher-level public facilities. Because detailed information on the health choice behavior of individuals as well as on all potential sources of supply is needed to understand bypassing, there are few empirical studies of this phenomenon. Analyses document large amounts of bypassing in Sri Lanka, where detailed surveys of health care supply and demand were collected (Akin and Hutchinson 1999; Samrasinghe and Akin 1994). Samrasinghe and Akin (1994) show that 31 percent of all episodes of illness were self-treated without medical consultation. Of those people who did seek treatment, only 42 percent were treated at the closest facility—either a predominantly private ayurvedic, or traditional, practice or a government-run, low-level health facility offering a Western approach to care; 58 percent did not go to the nearest facility. Of the non-ayurvedic sources, the public facilities (the type that primary health care would promote) are most frequently bypassed. Most individuals who bypass either a minor or a major public facility do so to visit a private Western facility.

The authors find that the bypassed public facilities had fewer doctors, nurses, and services and less equipment than public facilities that were not bypassed. In contrast, the private Western facilities that were bypassed had more doctors, nurses, and services and higher levels of care. Although this result might seem paradoxical, it is consistent with sophisticated health-seeking behavior. Because prices tend to be lower in the public facilities, patients will bypass high-quality, but expensive, private facilities in favor of public ones when their condition is not serious or quality is not important. For serious conditions, however, or when quality of service is important, individuals are willing to pay—in terms of both time and fees—for higher-quality care.

Bypassing can lead to low utilization of available public facilities. A survey (PIEDR 1994:vi) of a rural area of Punjab Province, Pakistan, found that although the physical infrastructure of rural primary health care was in place, "[o]nly about 5 percent of the sick children were taken for treatment to primary health care facilities; half were taken to private dispensers, and another quarter to private . . . doctors. Around 95

percent of deliveries took place at home." Roughly the same percentage of respondents sought treatment from a public rural health facility (5.2 percent) as from a "quack" (4.9 percent) (PIEDR 1994:35). The decision to bypass the public facility did not reflect long queues at those facilities. On the contrary, the typical rural health center saw only about 30 patients a day and the typical basic health center only 11 patients a day—far below their usual capacity, as rural health centers employed an average of 8 workers and basic health units, 5. Two atypically busy rural health centers attracted and serviced an average of more than 450 patients a day.

A study of health centers in Indonesia also found low usage of public facilities. Annual caseloads were low even for facilities located near large local populations (World Bank 1994b). On the basis of detailed case studies, this study identified two principal reasons for the low usage. First, many public facilities were short of equipment, drugs, and appropriate health workers. Second, and more important, detailed assessments on the way public health clinics operated showed that poorly functioning facilities contributed to patients' decisions to seek medical care elsewhere. Respondents in one case study made it clear that public facilities were of low quality. "They were confident that they could get considerate and unrushed care in a pleasant and informal setting in the private practice of doctors, bidans [midwives] and nurses" (World Bank 1994b:10).

Quantitative and focus group techniques in El Salvador found remarkably similar results (Lewis, Eskeland, and Traa-Valerezo 1999:26–27). Respondents consistently complained about the low quality of public health units, especially compared with the services provided by health centers and hospitals. Typical complaints about the public health units were: "Health posts operate only twice a week. Consultation is only until noon. The doctor is not always there. Sometimes only the nurse assistant is present. Waiting time is three hours on average. Only those who arrive by 8 get a consultation." In one locality, Potrero Sula, a respondent is quoted as saying, "The [public health] post here is useless because there is no doctor or nurse, and it is only open two days a week until noon." In contrast, for example, a respondent in El Pinar described the private health center at La Palma as "a little hospital with very good services. It is well equipped. The fee is only 3 colones for consultation and sometimes medication."

The study in El Salvador assessed the effect of intervention at the very lowest level—that of the health worker who lives in the community and provides simple medical services but refers seriously ill patients to larger facilities. Focus group respondents had very little use for these public health workers, and regression analysis found that they had little or no demonstrable effect on individuals' decisions about where to seek medical treatment.

Even when medical care in the public sector is of reasonable quality, it may be tremendously inefficient. A stark example of the kinds of inefficiency possible in the public sector comes from a detailed study of costs and expenditures in a public hos-

pital in the Dominican Republic, measured through careful observations on use of time (Lewis, La Forgia, and Sulvetta 1996). The study showed that although spending on personnel constituted 84 percent of total recurrent spending, actual staff costs for treating patients were only 19 percent of total costs. Gross inefficiency was identified as the cause of this huge discrepancy, an inefficiency that could be explained by the incentives, or lack thereof, facing personnel. There was no accountability for physicians or nurses, no rewards for extraordinary performance, and no punishment for inadequate or nonexistent performance. Salaries were low and undifferentiated. There was no management control over staff and essentially no returns to effective management. Clearly, the private sector also has problems, but the profit motive acts as a powerful incentive for efficiency.

How Can the Public Sector Deliver High-Quality Services?

Even strong advocates of primary health care would agree that governments do not always provide high-quality health services cost-effectively. Conversely, even the most ardent critic of government acknowledges that there are admirable and well-functioning health facilities and agencies in the public sector. Is failure, then, the result of resource constraints, ignorance, and mistakes, or is it a systematic and expected result of (dis)incentives created by institutional and organizational arrangements? If one is convinced that existing public sector problems are easily remediable through larger budgets, earmarked inputs, additional training, or technical assistance of various kinds, there is no reason to back away from public provision as a strategy. But if one is convinced that the failures are endemic and intrinsic to the especially low capability of the public sector, then perhaps the entire strategy for delivering health services needs rethinking.

PAY, EMPLOYMENT, AND PERFORMANCE IN THE HEALTH SECTOR. There is no one right answer about capability in the public health sector. Some public agencies provide high-quality, cost-effective health care. Others are capable of spending unlimited amounts of resources with no health gains. Two tough policy questions must be addressed. First, are the conditions in place for effective public provision of particular health services? And second, if not, is it possible to establish those conditions within a reasonable time horizon? Although some critics have perhaps been overly pessimistic about government capacity, supporters of publicly provided clinical services have paid far too little attention to the first question and have been entirely too sanguine about the second.

The feasibility of linking pay and employment to work effort in the public sector is not a new issue. It is generally recognized that the more essential and the less easily observable is individual effort, the greater is the importance of linking pay and performance (Milgrom and Roberts 1992). Workers in situations where effort is easily

observed and monitored are usually paid wages or salaries, whereas those whose output is observed but whose effort is not, such as salesmen, are usually paid on the basis of outcomes. In addition to the level of pay, there is the question of continued employment. Where output is crucially linked to individual performance and there can be little tolerance for deviations from high quality, continued employment is generally linked to performance. Observation of pay and employment across the public and private sectors tends to reinforce this position. To illustrate, table 1 shows a matrix of jobs and the degree to which pay and employment tend to be related to performance.

Among private sector professionals (lawyers, say, or doctors in medical practices), both pay and employment tend to be highly related to performance; in traditional public sector organizations, neither is. Such an observation is perhaps a commonplace. Yet rarely have policymakers or economists questioned whether the public sector is the most appropriate provider of clinical services, given that outsiders can observe neither health efficacy nor client treatment.

One reason for this lack of inquiry is that sometimes the most obvious and seamlessly working features of a system are invisible and are taken for granted when the system is functioning well. As a consequence, when one attempts to extrapolate from one set of social, legal, and political conditions to another, very different, set, the key features may be missed. For example, why will local health workers do the right thing even though there is no disciplining device of consumer choice? They are underpaid, and there are no effective institutional controls or legal restraints (for example, malpractice suits). The common (implicit) answer is often something like "because they are trained health professionals." Indeed, in well-functioning systems this is most likely the right answer. Doctors and nurses do not perceive themselves as performing for the money, or because of threats or close scrutiny, but out of professional pride and affiliation. The underlying factors of compensation and punishment are invisible not because they are weak or absent but precisely because they are so strong and effective that gross deviations from appropriate behavior are rare, and hence the need

Table 1. *Various Types of Employment and Compensation Schemes*

Degree to which employment is linked to performance	*Degree to which pay is linked to performance*		
	Low	*Medium*	*High*
Low	Traditional civil service arrangements (postal workers, administrators)	Large, stable organizations	Piece rate (agricultural workers, salesmen, contract workers)
High	"Up or out" organizations (U.S military)	Most private sector organizations	Professionals (law firms, medical practices)

Source Authors' summary

to invoke explicit punishment is similarly rare. When the underlying control mechanisms are weak or nonexistent, however, professionalism is not a powerful enough inducement.

The second incentive issue involves the proper mix of public and private providers. Such a mix is advantageous because mobility between the two sectors can act as a disciplining device on the behavior of individuals in the public sector. Doctors working in a public hospital when there is a large and effective private sector do not want to damage their reputations by performing noticeably worse than their colleagues. With either too little or too much interchange, however, this regulating mechanism will cease to be effective. The former effect may be due to a small private sector, the latter to a large private sector with professionals working in both sectors and with no standards or monitoring of the public part of their practice.

LEARNING FROM SUCCESS. What are the lessons from those instances in which the public sector is effective at improving health outcomes? Close examination points to the importance of social, political, and institutional factors in motivating effective performance from health workers. However, although the declaration at Alma Ata stated that one manifestation of these factors, "community participation," was a key feature of the primary health care strategy, it appears to be more the exception than the rule.

The role of social and political factors in generating the effective performance of government agencies in Kerala is well described in Heller (1996). Earlier, Caldwell (1986) described instances in which Keralans held health workers accountable through strong-armed means: "Doctors and others who provide village services (for instance, bus drivers plying regular routes) know stories of their fellows who were treated violently or hurt in protests about their having failed their duty" (p. 199). He quotes from a colleague working in West Bengal, where success was achieved "because [the state's communist government] appointed cadres at every health center to report on doctors or nurses who did not put all their time and effort into their services or who discriminated between patients" (p. 203).

A second example comes from an assessment of a major health intervention in Ceara State, northeastern Brazil, that contributed to a 36 percent fall in infant mortality in only a few years. Tendler and Freedheim (1994:1777–78) identify three primary reasons behind this success. First, the state used a merit hiring system and a large advertising campaign to create "a sense of 'mission' around the program and remarkable respect for its workers in the communities in which they served." Second, flexibility in job descriptions allowed workers to take on tasks that, although "sometimes viewed as distractions by experts, formed the basis for relations of trust between workers and citizens." Third, job candidates who were rejected were educated about what to expect from workers, supervisors, and elected officials, turning them into "informed public monitors of a new program in which the potential for

abuse was high." For example, job applicants were told, "Those of you who are not selected must make sure that those who are chosen abide by the rules. . . . If these rules are breached we want to hear about it. . . . [W]e are keeping all the applications, just in case any of those we hire do not perform well."

Motivating public (and private) sector workers to deliver high-quality services is a long-standing issue (and not just in the health sector), especially in situations where monitoring performance is difficult or costly. In the examples here, such motivation seems to have been achieved through direct monitoring with the threat of job loss, community monitoring with various threats, and community oversight and participation with the threat of job (and prestige) loss.

Public Sector Provision and the Total Use of Services

Why does the evidence show that even large-scale provision of some types of primary health care services has little or no impact on health status? One reason is that extending publicly funded health care could merely crowd out the consumption of equally effective private services. Even if the government were to deliver health services effectively, the health impact would depend not on the total use of public services but rather on how public provision affected total use of all services.

Although one might think at first blush that overall service use would increase by the same amount as the increase in government service capacity, the net effect of increasing public supply is always less than one for one. Only if there were no private sector or if there were no substitution between public and private services would total service increase by the same amount as the increase in government service. The degree to which public spending affects overall health services depends on four factors[6]

- *The smaller the overall elasticity of demand for health services, the smaller the impact of public spending.* Low elasticity of demand (demand that is unresponsive to price and distance) may well be the crux of the matter. Economists tend to shy away from discussions of "need" because it is not directly observable and is an emotionally charged term. But if an individual "needs" something, she or he is likely to forgo other consumption in order to satisfy that need. If income falls or prices rise, consumption of such a good will be protected, and the adjustment will be greater with regard to other uses of money and time. Because it is assumed that individuals place a high value on necessary health services (or at the very least those that alleviate unpleasant symptoms), the demand for them would be fairly insensitive to circumstances such as price or availability. Therefore, for those health services that are a matter of life or death, low price elasticities of demand should be expected. To the extent that severe symptoms are associated with health problems that are best treated in a clinic (which is not always the case),

these are the problems that are likely to be treated more or less successfully, regardless of a public presence. Makinen and Raney (1994:17) summarize five studies of medical care demand and find consistently that "price influences choice among providers and does so more strongly than its influence on whether or not to use services at all."

Indeed, in China (Cretin and others 1990), Indonesia (Gertler and Molyneaux 1995), and the United States (Manning and others 1987; Newhouse 1995), studies confirm that the demand for treatment is less elastic for serious conditions than for less serious conditions. Market failures associated with the treatment of serious health problems are likely to manifest themselves in ways other than bad health. For example, people may be overcharged, service may be inconvenient, or assets may be liquidated at inopportune moments (to get ready cash). All of these may result in inefficiency. But the impact on health, especially mortality, will be small.

The degree to which people take care of their health problems depends on the time elapsing between the onset of symptoms and treatment. In emergency situations, to take an extreme case, the effectiveness of treatment is tightly time-bound, allowing no opportunity to shop around The demand for treatment will be inelastic to its effective price. However, other dimensions of access that act as barriers to emergency care, such as emergency transport or credit, might be amenable to government interventions.

Health problems that can be treated if detected early but that have symptoms which are relatively minor at early stages may generate a demand for examinations that is highly elastic with respect to price or accessibility. In such cases a different dynamic provides another way in which government subsidies may well increase the socially optimal use of services. Screening for hypertension and some cancers falls into this category. This line of reasoning relies on subtle interactions of the severity of symptoms, the effectiveness of treatment, and, most important, the responsiveness to price and distance of the demand for care. In general, however, one would expect more serious illnesses to be associated with more inelastic demand for public services and, therefore, less potential influence of government policy on health status.

- *The larger the private sector, the smaller the impact of public spending.* According to the *World Development Indicators* (World Bank 1999), between 1990 and 1997 private (nongovernment) spending made up the largest percentage of health spending in nearly all poor countries, accounting for 59 percent in low-income countries (those with gross national product per capita of $785 or less in 1997) and almost 75 percent in South Asia (table 2).

Direct evidence on visits and treatments is hard to come by. A study in five Indian states, summarized in table 3, shows that in rural areas 82 percent of treated illness episodes were treated by private providers and that expenditures

Table 2. *Share of Total Health Expenditures That Is Private*

Category	*Average (percent)*	*Number of countries*
All countries	50.9	123
Income group		
Low income	58.7	34
Lower middle income	40.2	34
Higher middle income	47.9	24
High income, non-OECD	49.1	8
High income, OECD	33.3	23
Regions		
East Asia and Pacific	47.2	14
Latin America and the Caribbean	58.5	35
Middle East and North Africa	48.8	12
South Asia	74 9	5
Sub-Saharan Africa	51 7	22
Other	28.5	35

Notes Values are population-weighted means of country averages for the 1990–97 period OECD, Organisation for Economic Co-operation and Development

Source Derived from the World Development Indicators 1999 database.

for nonhospital treatment were nearly two-thirds (65 percent) of all out-of-pocket expenditures (World Bank 1995a). Note that private providers receive a larger part of these nonhospital expenditures (86 percent) than of expenditures as a whole because people can pay out of pocket for relatively minor illnesses but are less able to cover hospital expenses.

- *The greater the extent to which people see the private sector as a substitute for the public sector, the smaller the impact of public spending.* In addition to the size of the private sector, there is mounting evidence on the substitutability between public and private health care providers. Users may substitute private for public providers on the basis of price, location of facility, or quality of services. Several studies on the effect of increases in fees at (or distance to) public clinics look at the probability that a sick individual will attend a public versus a private clinic (table 4). Although results are not the same across all countries, both sets of characteristics have an effect on demand for care from private providers. The "cross-price" effect—for example, the impact of a change in fees at private clinics on the use of public clinics—is not negligible, especially if compared with "own-price" effects. For example, in Nigeria 100 percent of those who are deterred by higher prices in the public sector sought care from the private sector; in Ghana 60 percent of those deterred went to private providers; in El Salvador the fraction was 50 percent. Clearly, there is a high degree of substitutability between the two types of providers. Similarly, construction of new public facilities closer to consumers (reducing travel distance) can also displace demand from private

Table 3. *Demand for Private Sector Health Care in Five Indian States*

Area	*Gujarat*	*Maharashtra*	*Tamil Nadu*	*Uttar Pradesh*	*West Bengal*	*Weighted average*
Rural						
Percentage of treated illness episodes going to private providers	69	78	71	91	83	82
Percentage of out-of-pocket spending going to nonhospital treatment	62	64	74	59	74	65
Percentage of out-of-pocket spending going to private nonhospital treatment	48	55	68	51	65	56
Percentage of out-of-pocket spending on nonhospital treatment going to private treatment	77	86	92	86	88	86
Urban						
Percentage of treated illness episodes going to private providers	82	75	69	85	78	79
Percentage of out-of-pocket spending going to nonhospital treatment	61	60	69	57	63	61
Percentage of out-of-pocket spending going to private nonhospital treatment	53	51	62	46	54	52
Percentage of out-of-pocket spending on nonhospital treatment going to private treatment	87	85	90	81	86	85

Source Adapted from World Bank (1995a)

providers in the vicinity of the new facility. This effect is estimated to be substantial in Ghana and Kenya.

Another factor influencing the degree of substitutability between public and private providers is the absolute (expected) expense of the treatment sought. As discussed above, patients bypass both public and private facilities depending on their circumstances. Without insurance, people cannot afford to pay for expensive treatments when hospital care is essential; there is no plausible alternative to a public facility. For relatively cheap treatments, a private sector can thrive even without health insurance, and it is for such treatments that substitution between the sectors will be high.

- *The larger the private sector response to public intervention, the smaller the impact of public spending.* Changes in the price or availability of government interventions may induce a private supply response that can mitigate any actual impact on health outcomes. A case-control experiment in Indonesia showed that increasing user fees in public facilities caused a decline in use of those facilities and an expansion of the private health care sector; as a result, the effect of a user fee

Table 4. *Studies on the Probability of Provider Chosen: Percentage Deterred from Public Facilities Who Go to Private Facilities*

Sample	Change	Percent
Increase in price		
Benin (rural)	Community health center fees	61
Bolivia (urban)	Fees in Ministry of Public Health facilities, effect on adults 16 and over	33
	Fees in Ministry of Public Health facilities, effect on children 15 and under	0
	Fees in Ministry of Public Health facilities, effect on children 4 and under	0
El Salvador (urban)	Fees in Ministry of Health facilities, effect on males	56
	Fees in Ministry of Health facilities, effect on females	50
Ghana	Public fees	60
Kenya (rural)	Fees in government facilities from 0 to 10 Kenyan shillings, effect on adults 15 and over	37
Nigeria	Public prices, effect on adults 16 and over	100
Pakistan (urban)	Government clinic price, effect on children 5 and under	71
Increase in distance or time		
Bolivia (urban)	Waiting time in Ministry of Public Health facilities, effect on adults 16 and over	100
	Travel time to Ministry of Public Health facilities, effect on children 15 and under	20
	Waiting time in Ministry of Public Health facilities, effect on children 4 and under	20
Ghana	Distance to nearest public facility	50
Kenya (rural)	Distance to government facilities	44

Note See Filmer, Pritchett and Hammer (1998) for more details on the results in the underlying studies

Sources Derived from the following sources Benin—Bolduc, Lacroix, and Muller (1996), Bolivia—Ii (1996) El Salvador—Bitran and McInnes (1993), Ghana—Lavy and Germain (1994), Kenya—Mwabu, Ainsworth and Nyamete (1993), Nigeria—Akin, Guilkey, and Denton (1995), Pakistan—Alderman and Gertler (1989)

increase on overall health outcomes was very small (Gertler and Molyneaux 1995). In Tanzania, after the provision of health care was liberalized in 1991, the number of private sector providers skyrocketed; the number of nonprofit units operated by "approved voluntary organizations" increased from 697 to 780, and the number of private, for-profit facilities increased from 41 to 1,340 (Munishi 1997). The private sector expansion was dominated numerically by lower-level facilities such as dispensaries (from 36 to 1,313), but the number of private, for-profit hospitals also increased, from 4 to 20, while the number of government hospitals remained constant at 77.

A review of public spending in 13 regions in the Philippines (World Bank 1995b) from 1983 to 1990 found that public spending on health improved infant mortality in regions with lower incomes but had virtually no effect in richer regions. This result is consistent with substitutability of the private and public sectors in higher-income regions, with a resulting highly elastic private supply (and small impact on outcomes) and a less elastic supply response in poorer regions, with a substantial impact of public spending (Hammer 1997).

Although analyzing the potential impact of public sector interventions on the private sector might seem an obvious undertaking, it often has been ignored. A recent internal review of the World Bank—an institution with a large number of economists—found that of 217 appraisals of Bank health projects since 1970, *only 4* (all of these since 1995) had developed hypotheses regarding (or even considering) potential crowding-out (or crowding-in) as a result of the public supply of health inputs (World Bank 1998b).

Conclusion

Why is the evidence so thin for the seemingly plausible argument that primary health care is the best way to improve health in developing countries? We have argued that there are two weak links in the chain between providing public money for primary health care activities and achieving better health outcomes. The first is that governments often find it difficult to translate public spending into effective services, and the second is that of all the types of health services that the public sector might provide, primary care is the one that the private sector is likely to move away from as a result of any increase in public supply. Neither of these points is usually considered when discussing the applicability of a primary health care paradigm to a particular country.

What can be done? An alternative way of looking at policy in the health sector focuses on a careful analysis of the characteristic market failures in a country's health sector, the degree to which policy benefits reach the poor, and the government's ability to provide various kinds of services. Improvement of the quality of government health services depends on pressures outside the bureaucracy—citizen voice and consumer choice. From this perspective, clear recommendations can follow. First, public health activities should emphasize control of infectious diseases, largely through environmental changes. Second, public health programs should include demand-side mechanisms to improve routine clinical care. And third, policymakers should reconsider the role of hospital services as a practical means for governments to address inadequacies in insurance markets. The details of these arguments are available in a companion paper, "A Prescription for Health Policy in Poor Countries" (available from the authors).

Notes

Deon Filmer is an economist with the World Bank's Development Research Group, Jeffrey S. Hammer is lead economist with the Development Research Group, and Lant H. Pritchett is a lecturer at the John F. Kennedy School of Government, Harvard University, on leave from the World Bank. This paper grew out of an earlier collaboration between the authors and Maureen Lewis and Samuel Lieberman. The authors thank Philip Musgrove, Martin Ravallion, and Susan Stout for helpful discussions and comments.

1. This statistic is based on a comparison of average mortality rates for children under five in low- and middle-income countries with rates in high-income countries (World Bank 1999; see also Gwatkin 1980).

2. An entirely different viewpoint, also common in the literature, is that medical care of any sort—public or private—has had little, if any, impact on overall health measures. In this view, the large declines in mortality in the West preceded the development of medical treatments effective in combating its major causes. Better nutrition and housing conditions resulting from higher income as well as better hygienic practices are more likely responsible. We do not dispute this interpretation and note that it argues even more strongly for getting the first link right—making sure the composition of spending is geared toward basic hygiene—safe water, sanitation, and education.

3 Along with nearly all of the rest of the literature, we assume a fixed budget and avoid the problem of maximizing welfare, as that would endogenize the budget and require valuing health versus nonhealth goods.

4 In addition, the huge effort to create measures of disability-adjusted life years (DALYs) has led to some additional information on morbidity (Murray 1994) However, the correlation between DALYs lost and life expectancy or infant mortality is a high 0.93 across the eight regions for which DALYs have been calculated (World Bank 1993)

5 In that analysis, two-stage least-squares estimation was used to address the potential problems of measurement error and reverse causation. The estimates will be biased toward zero if these variables are measured with error (this is likely because a value for 1985 is sometimes used in place of one for 1990) In addition, the estimates will be inconsistent if there is reverse causation—if, for example high mortality rates induce the government to spend more on access to local health services

6 In other language, if D is total demand, S_b is public supply, ε_D is the elasticity of demand, ε_S is the elasticity of supply, and D_v is the demand for private services, then, allowing public and private services to be imperfect substitutes, with the elasticity of demand for private services with respect to public services $\varepsilon_{vb} < 0$, leads to $(dD/dS_b) = 1 + [(\varepsilon_S \times \varepsilon_{vb})/(\varepsilon_S - \varepsilon_D)] \times D_v/S_b$

References

The word "processed" refers to informally reproduced works that may not be commonly available through library systems

Agrawal, Amar Nath, Hari Om Varma, and R C. Gupta, eds 1996 *India Economic Information Year-Book* New Delhi: National

Akin, John S , and Paul Hutchinson. 1999. "Health-Care Facility Choice and the Phenomenon of Bypassing." *Health Policy and Planning* 14(2):135–51.

Akin, John S., David K. Guilkey, and E. Hazel Denton 1995. "Quality of Services and Demand for Health Care in Nigeria A Multinomial Probit Estimation." *Social Science and Medicine* 40(11) 1527–37

Alderman, Harold, and Paul Gertler. 1989. *The Substitutability of Public and Private Health Care for the Treatment of Children in Pakistan* Living Standards Measurement Study Working Paper 57 Washington, D.C.: World Bank.

Alderman, Harold, and Victor Lavy. 1996. "Household Responses to Public Health Services: Cost and Quality Tradeoffs." *The World Bank Research Observer* 11(1).3–22

Anand, Sudhir, and Martin Ravallion. 1993. "Human Development in Poor Countries: On the Role of Private Incomes and Public Services." *Journal of Economic Perspectives* 7(1):133–50.

Benefo, Kofi, and T. Paul Schultz. 1996. "Fertility and Child Mortality in Côte d'Ivoire and Ghana " *The World Bank Economic Review* 10(1):123–58.

Bidani, Benn, and Martin Ravallion. 1997. "Decomposing Social Indicators Using Distributional Data " *Journal of Econometrics* 77.125–39.

Bitran, Ricardo A., and D. Keith McInnes. 1993. *The Demand for Health Care in Latin America: Lessons from the Dominican Republic and El Salvador.* Economic Development Institute Seminar Paper 46. Washington, D.C.· World Bank.

Bolduc, Denis, Guy Lacroix, and Christophe Muller. 1996 "The Choice of Medical Providers in Rural Benin. A Comparison of Discrete Choice Models " *Journal of Health Economics* 15:477–98.

Caldwell, John C. 1986. "Routes to Low Mortality in Poor Countries " *Population and Development Review* 12(2):171–220.

Cretin, S., N. Duan, A Williams, X. Gu, and Y. Shi. 1990 "Modeling the Effect of Insurance on Health Expenditures in the People's Republic of China." *Health Services Research* 25(4):667–85.

DaVanzo, Julie. 1984. "A Household Survey of Child Mortality Determinants in Malaysia." *Population and Development Review* 10(Supplement):307–23

Decosas, Josef. 1990. "Planning for Primary Health Care: The Case of the Sierra Leone National Action Plan " *International Journal of Health Services* 20(1)·167–77

Filmer, Deon, and Lant Pritchett. 1999. "The Impact of Public Spending on Health: Does Money Matter?" *Social Science and Medicine* 49(1):1309–23.

Filmer, Deon, Lant Pritchett, and Jeffrey Hammer 1998 "Health Policy in Poor Countries: Weak Links in the Chain?" Policy Research Working Paper 1874. World Bank, Development Research Group, Poverty and Human Resources, Washington, D.C Processed.

Foster, S. D 1990 "Improving the Supply and Use of Essential Drugs in Sub-Saharan Africa." Policy Research Working Paper 456 World Bank, Population and Human Resources Department, Washington, D.C. Processed

Fox, Tamara 1995. "Primary Health Care· Notes on the Literature and Annotated Bibliography." Paper prepared for the World Bank, Latin America and Caribbean Human Development Group, Washington, D C Processed.

Frankenberg, Elizabeth. 1995 "The Effects of Access to Health Care on Infant Mortality in Indonesia " *Health Transition Review* 5.143–63

Gertler, Paul, and John Molyneaux. 1995 "Pricing Public Health Services: Lessons from a Social Experiment in Indonesia." RAND Corporation. Santa Monica, Calif Processed.

Griffin, Charles C 1992. *Health Care in Asia A Comparative Study of Cost and Financing.* Washington, D.C : World Bank.

Gwatkin, Davidson R. 1980. "How Many Die? A Set of Demographic Estimates of the Annual Number of Infant and Child Deaths in the World " *American Journal of Public Health* 70:1286–89.

Hammer, Jeffrey S 1997. "Economic Analysis for Health Projects." *The World Bank Research Observer* 12(1):47–71

Hammer, Jeffrey S., Ijaz Nabi, and James Cercone. 1995. "Distributional Effects of Social Sector Expenditures in Malaysia, 1974 to 1989." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence* Baltimore, Md.: Johns Hopkins University Press

Heller, Patrick. 1996. "Social Capital as a Product of Class Mobilization and State Intervention Industrial Workers in Kerala, India." *World Development* 24(6):1055–71.

Hossain, Shaikh. 1989. "Effect of Public Programs on Family Size and Child Education and Health" *Journal of Development Economics* 30(1):144–58.

Ii, Masako. 1996 *The Demand for Medical Care: Evidence from Urban Areas in Bolivia* Living Standards Measurement Study Working Paper 123 Washington, D.C.: World Bank.

International Institute for Population Studies 1995. *National Family Health Survey (MCH and Family Planning) India 1992–93* Bombay.

Koenig, Michael A., Vincent Fauveau, and Bogodan Wojtyniak. 1991. "Mortality Reductions from Health Interventions: The Case of Immunization in Bangladesh." *Population and Development Review* 17:87–104.

Kunstadter, Peter. 1985. "Remarks." In Scott B. Halstead, Julia A. Walsh, and Kenneth S Warren eds , *Good Health at Low Cost* New York: Rockefeller Foundation.

Lavy, Victor, and Jean-Marc Germain. 1994. *Quality and Cost in Health Care Choice in Developing Countries* Living Standards Measurement Study Working Paper 105. Washington, D.C.· World Bank.

Lavy, Victor, John Strauss, Duncan Thomas, and Philippe de Vreyer 1996 "Quality of Health Care Survival, and Health Outcomes in Ghana." *Journal of Health Economics* 15.333–57.

Lewis, Maureen, Gunnar S. Eskeland, and Ximena Traa-Valerezo. 1999. "Challenging El Salvador's Rural Health Care Strategy." Policy Research Working Paper 2164 World Bank, Washington D.C Processed

Lewis, Maureen A., Gerald M. La Forgia, and Margaret B Sulvetta 1996. "Measuring Public Hospital Costs: Empirical Evidence from the Dominican Republic " *Social Science and Medicine* 43(2) 221–34

Liu Yuanli, William C Hsiao, Qing Li, Xingzhu Liu, and Minghui Ren. 1995. "Transformation of China's Rural Health Care Financing." *Social Science and Medicine* 41(8):1085–93.

Macro International. 1997 *Brasil Pesquisa Nacional sobre Demographia e Saude, 1996* Demographic and Health Survey Program. Calverton, Md.

Makinen, Marty, and Laura Raney. 1994 "Role and Desirability of User Charges for Health Services." Draft Working Paper. Abt Associates, Bethesda, Md Processed.

Manning, Willard G., Joseph P Newhouse, Naihua Duan, Emmett B Keeler, Arleen Leibowitz, and M Susan Marquis. 1987 "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* 77(3) 251–77

Menken, Jane, and James Phillips 1990 "Population Change in a Rural Area of Bangladesh, 1967–87 " *The Annals* 510 87–101.

Milgrom, Paul, and John Roberts 1992 *Economics, Organization, and Management* Englewood Cliffs, N.J.. Prentice-Hall.

Mtemeli, D N 1994 "An Investigation into Why Women in Mutasa District Prefer to Deliver at Home and Not in Health Institutions " Provincial Medical Direction, Manicaland, Zimbabwe Processed.

Muhuri, Pradip K., and Samuel H Preston. 1991 "Effects of Family Composition on Mortality Differentials by Sex among Children in Matlab, Bangladesh." *Population and Development Review* 17:415–34.

Munishi, Gaspar K. 1997 "Private Health Sector Growth following Liberalization in Tanzania. Some Policy Considerations " International Health Policy Program Working Paper Washington, D C Processed.

Murray, Christopher. 1994. "Quantifying the Burden of Disease: The Technical Basis for Disability-Adjusted Life Years." *Bulletin of the World Health Organization* 72(3):429–45

Musgrove, Philip. 1996. *Public and Private Roles in Health Theory and Financing Patterns*. World Bank Discussion Paper 339 Washington, D.C

Mwabu, Germano, Martha Ainsworth, and Andrew Nyamete. 1993 "Quality of Medical Care and Choice of Medical Treatment in Kenya. An Empirical Analysis." *Journal of Human Resources* 28(4):838–62.

Newhouse, Joseph P. 1995. *Free for All* Cambridge, Mass. Harvard University Press.

Paes de Barros, Ricardo, D R. Da Costa, and Rosane Mendoça 1998. "Determinants of Infant Mortality in Brazil: Evidence from Census Data." Working Paper Instituto de Pesquisa Económica Aplicada, Brasilia. Processed

Panis, Constantijn W. A., and Lee A Lillard. 1994 "Health Inputs and Child Mortality in Malaysia" *Journal of Health Economics* 13(4):455–89

PIEDR (Pakistan Institute for Environment-Development Action) and Project Management Team. 1994. "The State of the Public Sector Primary Health Care Services, District Sheikhupura, Punjab, Pakistan." Bamako Initiative Technical Report Series 31 Bamako Initiative Management Unit, United Nations Children's Fund, New York. Processed

Pitt, Mark, Mark Rosenzweig, and Donna Gibbons 1993 "The Determinants and Consequences of the Placement of Government Programs in Indonesia" *The World Bank Economic Review* 7(3):319–48

Preston, Samuel H 1980 "Causes and Consequences of Mortality in Less Developed Countries during the Twentieth Century" In Richard Easterlin, ed, *Population and Economic Change in Developing Countries* Chicago, Ill University of Chicago Press for the National Bureau of Economic Research.

———. 1986 "Mortality and Development Revisited." *Population Bulletin of the United Nations* No 18-1985. Department of International Economic and Social Affairs, United Nations, New York

Ratzan, Scott C., Gary L. Filerman, and John W LeSar 2000 "Attaining Global Health. Challenges and Opportunities." *Population Bulletin* 55(1) 1–48

Rosenzweig, Mark, and T P Schultz 1982 "Child Mortality and Fertility in Colombia. Individual and Community Effects" *Health Policy and Education* 2 305–48

Rosenzweig, Mark, and Kenneth Wolpin 1982 "Government Interventions and Household Behavior in a Developing Country. Anticipating the Unanticipated Consequences of Social Programs" *Journal of Development Economics* 10 209–25

Samrasinghe, Daya, and John Akin 1994 "Health Strategy and Financing Study. Final Report" Prepared for Government of Sri Lanka, Ministry of Health and Women's Affairs, Colombo Processed.

Sastry, Narayan. 1995 "Community Characteristics, Individual and Household Attributes, and Child Survival in Brazil" RAND Labor and Population Program Working Paper Series 95-13. RAND Corporation, Santa Monica, Calif. Processed

Taylor, Carl E., and R D Singh n.d "The Narangwal Population Study. Integrated Health and Family Planning Services" Rural Health Research Center, Narangwal, India. Processed

Tendler, Judith, and Sara Freedheim 1994 "Trust in a Rent Seeking World. Health and Government Transformed in Northeastern Brazil." *World Development* 22(2) 1771–91

Thomas, Duncan, Victor Lavy, and John Strauss 1996 "Public Policy and Anthropometric Outcomes in Côte d'Ivoire" *Journal of Public Economics* 61:155–92.

WHO (World Health Organization). 1988. *Alma-Ata to the Year 2000: Reflections at the Midpoint* Geneva.

World Bank. 1993. *World Development Report 1993: Investing in Health*. New York: Oxford University Press.

———. 1994a. *Better Health in Africa: Lessons and Experiences Learned*. Washington, D.C.

———. 1994b. "Indonesia: Indonesia's Workforce: Issues and Options." Report 12835. Population and Human Resource Division, East Asia and Pacific Regional Office, Washington, D.C. Processed

———. 1995a. "India: Policy and Finance Strategies for Strengthening Primary Health Care Services." Report 13042-IN. Washington, D.C Processed.

———. 1995b. "Philippines· Public Expenditure Review." East Asia Department I, Washington, D.C Processed.

——— 1998a. *Reducing Poverty in India. Options for More Effective Public Services* World Bank Country Study. Washington, D.C

———. 1998b. "Lessons from Experience in Health, Nutrition, and Population." Sector and Thematic Evaluations Group, Operations Evaluation Department, Washington, D.C. Processed

———. 1999 *World Development Indicators*. Washington, D.C.

Social Capital: Implications for Development Theory, Research, and Policy

Michael Woolcock • Deepa Narayan

In the 1990s the concept of social capital—defined here as the norms and networks that enable people to act collectively—enjoyed a remarkable rise to prominence across all the social science disciplines. The authors trace the evolution of social capital research as it pertains to economic development and identify four distinct approaches the research has taken: communitarian, networks, institutional, and synergy. The evidence suggests that of the four, the synergy view, with its emphasis on incorporating different levels and dimensions of social capital and its recognition of the positive and negative outcomes that social capital can generate, has the greatest empirical support and lends itself best to comprehensive and coherent policy prescriptions. The authors argue that a significant virtue of the idea of and discourse on social capital is that it helps to bridge orthodox divides among scholars, practitioners, and policymakers.

What is social capital? How does it affect economic development? What are the implications for theory, research, and policy? These questions lie at the heart of recent attempts to make sense of the burgeoning literature on social capital and to ascertain its relationship to economic development. In this article we endeavor to answer each of these questions; in so doing, we provide an overview of the scholarship on social capital for those unfamiliar with the term as well as a sense of coherence and direction to those embarking on new empirical research and policy analysis in this rich field.

What Is Social Capital?

"It's not what you know, it's *who* you know." This common aphorism sums up much of the conventional wisdom regarding social capital. It is wisdom born of

The World Bank Research Observer, vol 15, no 2 (August 2000), pp 225–49

experience—that gaining membership to exclusive clubs requires inside contacts, that close competitions for jobs and contracts are usually won by those with friends in high places. When people fall on hard times, they know it is their friends and family who constitute the final safety net. Conscientious parents devote hours to the school board and to helping their children with homework, only too aware that a child's intelligence and motivation are not enough to ensure a bright future. Some of our happiest and most rewarding hours are spent talking with neighbors, sharing meals with friends, participating in religious gatherings, and volunteering for community projects.

Intuitively, then, the basic idea of social capital is that a person's family, friends, and associates constitute an important asset, one that can be called on in a crisis, enjoyed for its own sake, and leveraged for material gain. What is true for individuals, moreover, also holds for groups. Those communities endowed with a diverse stock of social networks and civic associations are in a stronger position to confront poverty and vulnerability (Moser 1996; Narayan 1995), resolve disputes (Schafft 1998; Varshney 2000), and take advantage of new opportunities (Isham 1999). Conversely, the *absence* of social ties can have an equally important impact. Office workers, for example, fear being left out of the loop on important decisions; ambitious professionals recognize that getting ahead in a new venture typically requires an active commitment to networking. A defining feature of being poor, moreover, is that one is not a member of—or may even be actively excluded from—certain social networks and institutions that could be used to secure good jobs and decent housing (Wilson 1987, 1996).

Intuition and everyday language also recognize an additional feature of social capital that it has costs as well as benefits, that social ties can be a liability as well as an asset. Most parents, for example, worry that their teenage children will fall in with the wrong crowd and that peer pressure and a strong desire for acceptance will induce them to take up harmful habits. Even close family members can overstay their welcome. At the institutional level, many countries and organizations have nepotism laws, in explicit recognition that personal connections can be used to discriminate unfairly, distort, and corrupt. Everyday language and life experience, in short, teach that the social ties individuals have can be both a blessing and a blight, while those they do *not* have can deny them access to key resources. These features of social capital are well documented by the empirical evidence and have important implications for economic development and poverty reduction.

These examples suggest a more formal definition: social capital refers to the norms and networks that enable people to act collectively. This simple definition serves a number of purposes. First, it focuses on the sources, rather than the consequences, of social capital (Portes 1998) while recognizing that important features of social capital, such as trust and reciprocity, are developed in an iterative process. Second, this definition permits the incorporation of different dimensions of social capital and

recognizes that communities can have access to more or less of them. The poor, for example, may have a close-knit and intensive stock of "bonding" social capital that they can leverage to "get by" (Briggs 1998; Holzmann and Jorgensen 1999), but they lack the more diffuse and extensive "bridging" social capital deployed by the nonpoor to "get ahead" (Barr 1998; Kozel and Parker 2000; Narayan 1999). Accordingly, such an approach allows the argument that it is different *combinations* of bonding and bridging social capital that are responsible for the range of outcomes observed above and incorporates a dynamic component in which optimal combinations of these dimensions change over time. Third, while this definition presents the community (rather than individuals, households, or the state) as the primary unit of analysis, it recognizes that individuals and households (as members of a given community) can nonetheless appropriate social capital and that the way communities themselves are structured turns in large part on their relationship with the state. Weak, hostile, or indifferent governments have a profoundly different effect on community life and development projects, for example, than do governments that respect civil liberties, uphold the rule of law, honor contracts, and resist corruption (Isham and Kaufmann 1999).

This conceptualization of the role of social relationships in development represents an important departure from earlier theoretical approaches and therefore has important implications for contemporary development research and policy. Until the 1990s the major theories of development held rather narrow, even contradictory, views about the role of social relationships in economic development and offered few constructive policy recommendations. In the 1950s and 1960s, for example, traditional social relationships and ways of life were viewed as impediments to development. When modernization theorists explained "the absence or failure of capitalism," Moore (1997:289) correctly notes, "the focus [was] on social relations as obstacles." As an influential United Nations (1951) document of the time put it, for development to proceed, "ancient philosophies have to be scrapped; old social institutions have to disintegrate; bonds of caste, creed and race have to burst; and large numbers of persons who cannot keep up with progress have to have their expectations of a comfortable life frustrated" (cited in Escobar 1995:3).

This view gave way in the 1970s to the arguments of dependency and world-systems theorists, who held that social relations among corporate and political elites were a primary mechanism of capitalist exploitation. The social characteristics of poor countries and communities were defined almost exclusively in terms of their relation to the means of production and the inherent antipathy between the interests of capital and labor. Little mention was made of the possibility (or desirability) of mutually beneficial relationships between workers and owners, of the tremendous variation in the degree of success recorded by developing countries, or of political strategies—other than revolution—by which the poor could improve their lot. At the same time, communitarian perspectives stressed the inherent beneficence and

self-sufficiency of local communities but underestimated the negative aspects of communal obligations, overestimated the virtues of isolationism and self-sufficiency, and neglected the importance of social relations in constructing effective and accountable formal institutions. For their part, neoclassical and public choice theorists—whose voices were the most influential in the 1980s and early 1990s—assigned no distinctive properties to social relations. These perspectives, which focused on the strategic choices of rational individuals interacting under various time, budgetary, and legal constraints, held that groups (including firms) existed primarily to lower the transaction costs of exchange; given undistorted market signals, the optimal size and combination of groups would duly emerge.

The major development theories, then, construed social relations as singularly burdensome, exploitative, liberating, or irrelevant. Reality, unfortunately, does not conform so neatly to these descriptions and their corresponding policy prescriptions. Events in the post–cold war era—from ethnic violence and civil war to financial crises and the acknowledgement of widespread corruption—demand a more sophisticated appraisal of the vices, virtues, and vicissitudes of the social dimension as it pertains to the wealth and poverty of nations (Woolcock forthcoming). The literature on social capital, in its broadest sense, represents a first approximation to the answer to this challenge. It is a literature to which all the social science disciplines have contributed, and it is beginning to generate a remarkable consensus regarding the role and importance of institutions and communities in development. Indeed, one of the primary benefits of the idea of social capital is that it allows scholars, policymakers, and practitioners from different disciplines to enjoy an unprecedented level of cooperation and dialogue (Brown 1998; Brown and Ashman 1996).

Four Perspectives on Social Capital and Economic Development

The letter and spirit of social capital have a long intellectual history in the social sciences (Platteau 1994; Woolcock 1998), but the sense in which the term is used today dates back more than 80 years to the writings of Lyda J. Hanifan, then the superintendent of schools in West Virginia. Explaining the importance of community participation in enhancing school performance, Hanifan (1916:130) invoked the concept of social capital, describing it as

> those tangible substances [that] count for most in the daily lives of people: namely good will, fellowship, sympathy, and social intercourse among the individuals and families who make up a social unit. . . . If [an individual comes] into contact with his neighbor, and they with other neighbors, there will be an accumulation of social capital, which may immediately satisfy his

> social needs and which may bear a social potentiality sufficient to the substantial improvement of living conditions in the whole community.

After Hanifan the idea of social capital disappeared for several decades but was reinvented in the 1950s by a team of Canadian urban sociologists (Seely, Sim, and Loosely 1956), in the 1960s by an exchange theorist (Homans 1961) and an urban scholar (Jacobs 1961), and in the 1970s by an economist (Loury 1977). None of these writers, interestingly, cited earlier work on the subject, but all used the same umbrella term to encapsulate the vitality and significance of community ties. The seminal research by Coleman (1987, 1988, 1990) on education and by Putnam (1993, 1995) on civic participation and institutional performance, however, has provided the inspiration for most of the current work, which has since coalesced around studies in nine primary fields: families and youth behavior; schooling and education; community life (virtual and civic); work and organizations; democracy and governance; collective action; public health and environment; crime and violence; and economic development.[1]

In this paper we are concerned with this final category and related work in political economy and new institutional economics. Research on social capital and economic development can be categorized into four distinct perspectives: the communitarian view, the networks view, the institutional view, and the synergy view.

The Communitarian View

The communitarian perspective equates social capital with such local organizations as clubs, associations, and civic groups. Communitarians, who look at the number and density of these groups in a given community, hold that social capital is inherently good, that more is better, and that its presence always has a positive effect on a community's welfare. This perspective has made important contributions to analyses of poverty by stressing the centrality of social ties in helping the poor manage risk and vulnerability. As Dordick (1997) notes, the poor have "something left to lose"—each other.

In their celebration of community and civil society, however, many enthusiasts of this view of social capital have ignored its important downside (Portes and Landolt 1996). For example, where communities or networks are isolated, parochial, or working at cross-purposes to society's collective interests (in ghettos, gangs, drug cartels, and so on), productive social capital is replaced by what Rubio (1997)—in discussing Colombia—calls *perverse* social capital, which greatly hinders development. Many benefits certainly are associated with being a member of a highly integrated community, but there are also significant costs, and for some, the costs may greatly outweigh the benefits. Consider, for instance, the bright girls who are taken out of village schools in India because of community expectations. The social networks underly-

ing organized crime syndicates in Latin America and Russia may generate large negative externalities for society in the form of lost lives, wasted resources, and pervasive uncertainty. The communitarian perspective also implicitly assumes that communities are homogenous entities that automatically include and benefit all members. But the extensive literature on caste inequality, ethnic exclusion, and gender discrimination—the bleak outcomes often produced and maintained by community pressures—suggests otherwise (Narayan and Shah 1999).

Evidence from the developing world demonstrates why merely having high levels of social solidarity or informal groups does not necessarily lead to economic prosperity. In Kenya a participatory poverty assessment recorded more than 200,000 community groups active in rural areas, but most were unconnected to outside resources and were unable to improve the lot of the poor (Narayan and Nyamwaya 1996) A World Bank (1989) report on Rwanda cited more than 3,000 registered cooperatives and farmers groups and an estimated 30,000 informal groups, yet these groups were unable to prevent one of history's most gruesome civil wars. In many Latin American countries, indigenous groups are often marked by high levels of social solidarity, but they remain excluded economically because they lack the resources and access to power that are necessary to shift the rules of the game in their favor (Narayan 1999). This is also the case in Haiti, where social capital, "rich at the local level," is employed by peasant groups to "meet labor requirements, gain access to land, protect clientship in the marketplace, promote mutual aid, assure protection from state authorities, and generally manage risk." Even so, these groups cannot overcome the crippling effects of colonialism, corruption, "geographical isolation, political exclusion, and social polarization" (all quotations from White and Smucker 1998:1–3).

The Networks View

A second perspective on social capital, which attempts to account for both its upside and its downside, stresses the importance of vertical as well as horizontal associations between people and of relations within and among such organizational entities as community groups and firms. Building on work by Granovetter (1973), it recognizes that strong intracommunity ties give families and communities a sense of identity and common purpose (Astone and others 1999). This view also stresses, however, that without weak intercommunity ties, such as those that cross various social divides based on religion, class, ethnicity, gender, and socioeconomic status, strong horizontal ties can become a basis for the pursuit of narrow sectarian interests. In the recent popular literature, the former has been called "bonding" and the latter "bridging" social capital (Gittell and Vidal 1998). Different combinations of these dimensions, it is argued, are responsible for the range of outcomes that can be attributed to social capital. This more nuanced perspective, which we call the *networks* view, re-

gards the tension between social capital's virtues and vices as a defining property, one that explains in part why scholars and policymakers have been so persistently ambivalent about its potential as a theoretical construct and policy instrument.

The networks view of social capital is closely associated with Burt (1992, 1997, 1998); Fafchamps and Minten (1999); Massey (1998); Massey and Espinosa (1997); Portes (1995, 1997, 1998); and Portes and Sensenbrenner (1993). It is characterized by two key propositions. First, social capital is a double-edged sword. It can provide a range of valuable services for community members, ranging from baby-sitting and house-minding to job referrals and emergency cash. But there are also costs in that those same ties can place considerable noneconomic claims on members' sense of obligation and commitment, with negative economic consequences. Group loyalties may be so strong that they isolate members from information about employment opportunities, foster a climate of ridicule toward efforts to study and work hard, or siphon off hard-won assets (say, to support recent immigrants from the home country). Portes and Sensenbrenner (1993) cite the case of prosperous Asian immigrants who anglicized their names in order to divest themselves of communal obligations to subsequent cohorts. Second, the sources of social capital need to be distinguished from the consequences derived from them. Imputing only desirable outcomes to social capital, or equating them with it, ignores the possibility that these outcomes may be attained at another group's expense, that given outcomes may be suboptimal, or that desirable outcomes attained today come at the price of significant costs tomorrow.

These results have given rise to the logical conclusion that both strong intracommunity ties and weak extracommunity networks are needed to avoid making tautological claims regarding the efficacy of social capital. (Without this distinction, for example, it could be argued that successful groups are distinguished by their dense community ties, failing to consider the possibility that the same ties could be *preventing* success in another otherwise similar group.) Accordingly, the networks view argues that communities can be characterized by their endowments of these two dimensions of social capital and that different combinations of these dimensions account for the range of outcomes associated with social capital (table 1).

Furthermore, as community members' welfare changes over time, so too does the optimal calculus of costs and benefits associated with particular combinations of bonds and bridges. Poor entrepreneurs, for example, initially dependent on their

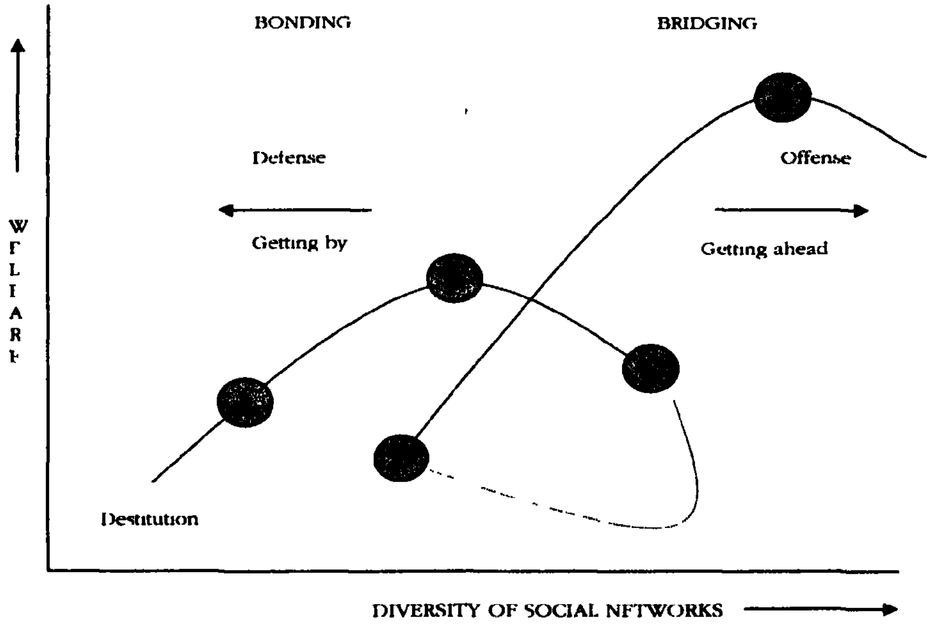
Table 1. *Dimensions of Social Capital at the Community Level*

Extracommunity networks (bridging)	*Intracommunity ties (bonding)*	
	Low	*High*
Low	Outcasts	Poor villagers
High	Recent rural-to-urban migrants	Successful members of microfinance programs

immediate neighbors and friends (their bonding social capital) for credit, insurance, and support, require access to more extensive product and factor markets as their businesses expand. Granovetter (1995) argues that economic development takes place through a mechanism that allows individuals to draw initially on the benefits of close community membership but that also enables them to acquire the skills and resources to participate in networks that transcend their community, thereby progressively joining the economic mainstream.

These insights can be demonstrated graphically and applied to poverty reduction more generally. Figure 1 shows that as the social networks of the poor become more diverse, so too does their welfare. The social capital residing in a given network can be leveraged or used more efficiently, which is essentially the genius of group-based credit programs such as the well-known Grameen Bank in Bangladesh (van Bastelaer 1999). Poor village women with no material collateral are given loans on the basis of their membership in a small peer group, which helps them start or expand a small business and thereby improve their families' welfare (A). But the economic returns

Figure 1. *Social Capital and Poverty Transitions*



Source Woolcock (2000)

o any given group soon reach a limit (B), especially when they rely on high endowments of "bonding" social capital. If the group continues to expand—for example, hrough the arrival of subsequent cohorts from the village—its resources may beome overwhelmed, thereby reducing the well-being of long-established members C). Similarly, long-term members of group-based credit programs may find that bligations and commitments to their colleagues present obstacles to further adancement, especially for the more ambitious (Woolcock 1999). In these circumtances, many poor people partially divest themselves of their immediate community ies (D) and find a potentially more diverse network where "bridging" social capital ; more abundant and economic opportunities more promising (E). Migration from illages to cities is the most dramatic example of this situation, but Portes and ensenbrenner's (1993) name-changing Asian immigrants are doing essentially the ame thing.

The networks view has been employed with great effect in recent development esearch. In their analysis of poor communities in rural areas of northern India, for xample, Kozel and Parker (2000) report that social groups among poor villagers serve vitally important protection, risk management, and solidarity functions. It is the more extensive and leveraged networks of the nonpoor, by contrast, that are used for strategic advantage and the advancement of material interests. Crudely put, the networks of the poor play defense, while those of the nonpoor play offense. Barr (1998) reports strikingly similar results from her work on the relationship between the structure of business networks and enterprise performance in Africa. Poor entrepreneurs, operating small local firms in traditional industries, form what Barr calls solidarity networks to exchange personal information about members' conduct and intentions. The primary function of these networks is to reduce risk and uncertainty. Larger regional firms, in contrast, coalesce into innovation networks that share knowledge about technology and global markets with the explicit goal of enhancing enterprise profit, productivity, and market share (see also Van Dijk and Rabellotti 1997; Fafchamps and Minten 1999). Far from dismissing the vitality of traditional village groups in poor communities (the modernization view) or romanticizing it (the communitarian view), the networks view in effect recognizes that these groups can both help and hinder economic advancement.

The clear challenge to social capital theory, research, and policy from the networks perspective is thus to identify the conditions under which the many positive aspects of bonding social capital in poor communities can be harnessed and its integrity retained (and, if necessary, its negative aspects dissipated), while simultaneously helping the poor gain access to formal institutions and a more diverse stock of bridging social capital. This process is fraught with multiple dilemmas, however, especially for external nongovernmental organizations, extension services, and development agencies, because it may entail altering social systems that are the product of longstanding cultural traditions or of powerful vested interests.

The particular strength of the networks view is its willingness to engage in detailed policy discussions on the basis of compelling empirical evidence and detailed assessments of the veracity of competing explanations. This view, however, minimizes the "public good" nature of social groups, regarding any benefits of group activity as primarily the property of the particular individuals involved. Its proponents thus are highly skeptical of arguments that social capital can (or should) be measured across larger social aggregates, such as societies or nations (Portes 1998). Neither does the networks approach explicitly incorporate institutions at the societal level and their capacity to both shape and be shaped by local communities. To be sure, the networks perspective recognizes that weak laws and overt discrimination can undermine efforts by poor minorities to act in their collective interest, but the role communities play in shaping institutional performance generally, and the enormous potential of positive state-society relations in particular, are largely ignored.

The Institutional View

A third perspective of social capital, which we call the *institutional* view, argues that the vitality of community networks and civil society is largely the product of the political, legal, and institutional environment. Where the communitarian and networks perspectives largely treat social capital as an independent variable giving rise to various outcomes, both good and bad, the institutional view instead views social capital as a dependent variable. This approach argues that the very capacity of social groups to act in their collective interest depends on the quality of the formal institutions under which they reside (North 1990). It also stresses that the performance of states and firms themselves depends on their own internal coherence, credibility, and competence and on their external accountability to civil society.

Research from the institutional view has two variants, both of which have yielded remarkably complementary results. The first approach, described by Skocpol (1995, 1996), encompasses case studies based on comparative history and contends that it is wrong to argue that firms and communities thrive to the extent that governments retreat. On the contrary, Skocpol shows, civil society thrives to the extent that the state actively encourages it. Tendler's (1997) research on the political economy of decentralization in Brazil similarly stresses the importance of good government for making local programs work.

A second, and increasingly influential, approach relies on quantitative cross-national studies of the effects of government performance and social divisions on economic performance. This approach, pioneered by Knack and Keefer (1995, 1997), equates social capital with the quality of a society's political, legal, and economic institutions. Drawing on various indexes of institutional quality compiled by investment agencies and human rights groups, these studies show that items such as "generalized trust," "rule of law," "civil liberties," and "bureaucratic quality" are positively

associated with economic growth. In a recent review of this particular strand of the literature, Knack (1999:28) concludes that "social capital reduces poverty rates and improves, or at a minimum does not worsen, income inequality."

Collier and Gunning (1999) employ a variation of this view in their analyses of the causes of slow growth in Africa (see also Collier 1998, 1999; Temple 1998). Distinguishing between civic and government social capital, they show that slow growth occurs in societies with both high levels of ethnic fragmentation and weak political rights. Although Rodrik (1998, 1999) does not employ the terminology of social capital, he makes a similar argument, demonstrating that economies with divided societies and weak institutions for managing conflict respond sluggishly to shocks. Easterly (2000) also reports that societies able to generate and sustain a middle-class consensus are those most likely to produce stable and positive rates of growth. The related literature on social capabilities and development (Hall and Jones 1999; Temple and Johnson 1998) tells a similar story.

Several empirical and methodological questions can be raised about these studies, but in aggregate their message is loud and clear. Rampant corruption, frustrating bureaucratic delays, suppressed civil liberties, vast inequality, divisive ethnic tensions, and failure to safeguard property rights (to the extent that they exist at all) are major impediments to prosperity. In countries where these conditions prevail, there is little to show for well-intentioned efforts to build schools, hospitals, roads, and communications infrastructure or to encourage foreign investment (World Bank 998). Investments in civic and government social capital are thus highly complementary to investments in more orthodox forms of capital accumulation.

The very strength of the institutional view in addressing macroeconomic policy oncerns, however, is also a weakness in that it lacks a microeconomic component. reedoms, rights, and liberties, for example, have to be secured by government. Coerent and competent bureaucracies may take decades to construct and may yield enefits more immediately suited to corporate interests than to those of the poor In roviding broad statistical evidence for the importance of social capital, the subtlety, chness, and enormous variation gleaned from case studies of individual countries nd communities is lost, as are the voices of those most directly affected by weak ublic institutions: the poor.

he Synergy View

n recognition of this disconnect, a number of scholars have recently proposed what ight be called a *synergy view*, which attempts to integrate the compelling work nerging from the networks and institutional camps. Although the synergy view aces its intellectual antecedents to earlier work in comparative political economy nd anthropology, its most influential body of research was published in a special sue of *World Development* (1996). The contributors to this volume examined cases

from Brazil, India, Mexico, the Republic of Korea, and Russia in search of the conditions that foster developmental synergies—dynamic professional alliances and relationships between and within state bureaucracies and various actors in civil society.

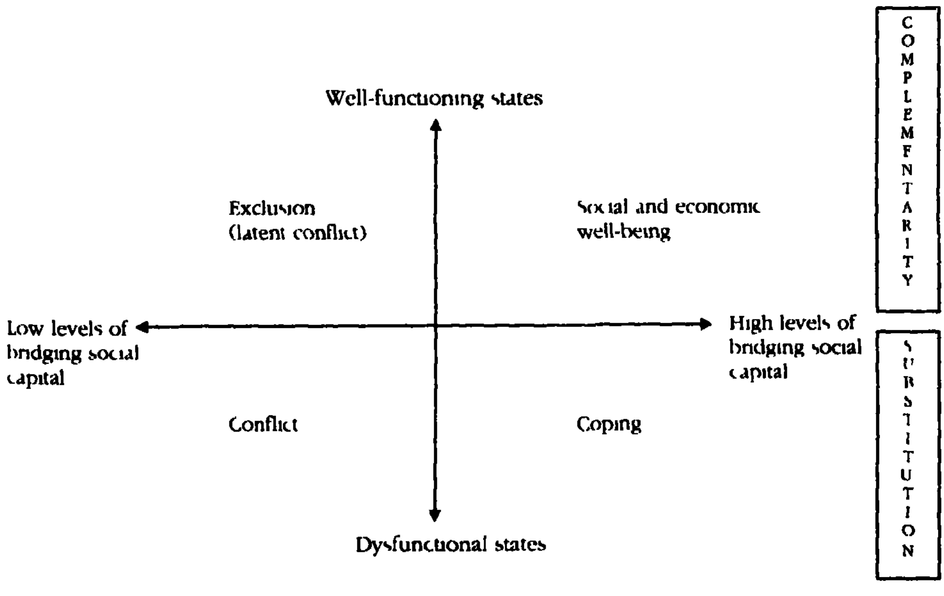
Three broad conclusions emerged from these studies:

- Neither the state nor societies are inherently good or bad; governments, corporations, and civic groups are variable in the impact they can have on the attainment of collective goals.
- States, firms, and communities alone do not possess the resources needed to promote broad-based, sustainable development; complementarities and partnerships forged both within and across these different sectors are required. Identifying the conditions under which these synergies emerge (or fail to emerge) is thus a central task of development research and practice.
- Of these different sectors, the state's role in facilitating positive developmental outcomes is the most important and problematic. This is so because the state is not only the ultimate provider of public goods (stable currencies, public health, universal education) and the final arbiter and enforcer of the rule of law (property rights, due process, freedom of speech and association) but is also the actor best able to facilitate enduring alliances across the boundaries of class, ethnicity, race, gender, politics, and religion. Communities and firms also have an important role to play in creating the conditions that produce, recognize, and reward good governance. In otherwise difficult institutional environments, community leaders who are able to identify and engage what Fox (1992) calls "pockets of efficiency within the state" become agents of more general reform.

Evans (1992, 1995, 1996), one of the primary contributors to this view, concludes that synergy between government and citizen action is based on *complementarity* and *embeddedness*. Complementarity refers to mutually supportive relations between public and private actors and is exemplified in legal frameworks that protect rights of association and in more humble measures such as chambers of commerce to facilitate exchanges among community associations and business groups. Embeddedness refers to the nature and extent of the ties connecting citizens and public officials. The classic examples are from irrigation, in which the lowest-level irrigation officials are from the community being served; they are enmeshed in local social relations and hence are under pressure by the community to perform and be responsive to them. Importantly, this approach works only where the actions of public officials are simultaneously bound by performance-oriented organizational environments that are competent, coherent, and credible. As the case of Russia amply demonstrates, weak public institutions and deep cleavages between powerful authorities and ordinary citizens can lead to political instability, rampant corruption, rising inequality, and capital flight (Rose 1998).

Developing these ideas, Woolcock (1998) shows that a range of development outcomes flows from different types and combinations of community capacity and state functioning. Narayan (1999) integrates the core ideas of bridging social capital and state-society relations and suggests that different interventions are needed for different combinations of governance and bridging social capital in a group, community, or society (figure 2). In societies (or communities) with good governance and high levels of bridging social capital, there is complementarity between state and society, and economic prosperity and social order are likely. But when a society's social capital inheres mainly in primary social groups disconnected from one another, the more powerful groups dominate the state, to the exclusion of other groups. Such societies, which include countries in Latin America with large excluded indigenous populations, are characterized by latent conflict. In these circumstances, a key task for subordinate groups and activists is to forge broad, coherent coalitions (Keck and Sikkink 1998) and nurture relations with allies in positions of power (Fox and Brown 1998); should they be successful, weak groups may begin to accrue rights and resources previously denied them. Similarly, a state that opens up and explicitly builds bridges

Figure 2. *Relationship between Bridging Social Capital and Governance*



Note Complementarity refers to the optimal interaction of government and markets in civil society, substitution is the replacement by informal organizations (families, networks, and so on) of services ordinarily provided by governments and institutions

Source Adapted from Narayan (1999)

to excluded groups increases the likelihood that the poor will be able to gain access to the resources and services to which they are entitled.

Alternatively, state-society relations may degenerate into conflict, violence, war or anarchy—a breakdown that allows warlords, local mafias, and guerrilla movements to take over the power and authority of the state. Restoring economic prosperity and peace to Rwanda, for example, will involve forging a measure of reconciliation between two ethnic groups. Often, when citizens are deprived of services and benefits, informal networks substitute for the failed state and form the basis of coping strategies. This is the case in Benin and Togo, where women, denied access to formal credit, established informal revolving credit societies; in Tanzania the absence of police protection has led some villages to rely on their own system of security guards (Narayan and others 2000).

When representatives of the state, the corporate sector, and civil society establish common forums through which they can pursue common goals, development can proceed. In these circumstances social capital has a role as a mediating variable that is shaped by public and private institutions. This shaping is an inherently contentious and political process, one in which the role of the state is crucial. Moreover, the fundamental social transformation of economic development—from traditional kinship-based community life to societies organized by formal institutions—alters the calculus of costs and benefits associated with different dimensions of social capital and the desirable combinations of these dimensions (Berry 1993). Although development struggles are inherently political, they are not always won by the most powerful, nor do challenges to authority always entail violent conflict. Patient efforts by intermediaries to establish partnerships between associations of the poor and outsiders can reap significant dividends (Isham, Narayan, and Pritchett 1995). As Uphoff (1992:273) points out,

> paradoxical though it may seem, "top-down" efforts are usually needed to introduce, sustain, and institutionalize "bottom-up" development. We are commonly constrained to think in "either-or" terms—the more of one the less of the other—when both are needed in a positive-sum way to achieve our purposes.

The synergy view suggests three central tasks for theorists, researchers, and policymakers: to identify the nature and extent of a community's social relationships and formal institutions, and the interaction between them; to develop institutional strategies based on these social relations, particularly the extent of bonding and bridging social capital; and to determine how the positive manifestations of social capital—cooperation, trust, and institutional efficiency—can offset sectarianism, isolationism, and corruption. Put another way, the challenge is to transform situations where a community's social capital substitutes for weak, hostile, or indifferent formal institutions into ones in which both realms complement one another.

Table 2. *Four Views of Social Capital*

Perspective	*Actors*	*Policy prescriptions*
Communitarian view		
Local associations	Community groups Voluntary organizations	Small is beautiful Recognize social assets of the poor
Networks view		
Bonding and bridging community ties	Entrepreneurs Business groups Information brokers	Decentralize Create enterprise zones Bridge social divides
Institutional view		
Political and legal institutions	Private and public sectors	Grant civil and political liberties Institute transparency, accountability
Synergy view		
Community networks and state-society relations	Community groups, civil society, firms, states	Coproduction, complementarity Participation, linkages Enhance capacity and scale of local organizations

Table 2 summarizes the key elements of the four perspectives on social capital and development and their corresponding policy prescriptions. The differences between them are primarily the unit of analysis on which they focus; their treatment of social capital as an independent, dependent, or mediating variable; and the extent to which they incorporate a theory of the state. The largest and most influential bodies of work have emerged from the networks and institutional perspectives; the most recent approaches seek a synthesis.

Measuring Social Capital

Several recent innovative studies have attempted to quantify social capital and its contribution to economic development. To arrive at concrete policy recommendations for using social capital as a development tool, more comparative research is required that uses precise measures of social capital to examine within-country and across-country variations in poverty reduction, government performance, ethnic conflict, and economic growth. Obtaining a single, true measure of social capital is probably not possible, for several reasons. First, the most comprehensive definitions of social capital are multidimensional, incorporating different levels and units of analysis. Second, the nature and forms of social capital change over time, as the balance shifts between informal organizations and formal institutions. And third, because no long-standing cross-country surveys were initially designed to measure social capital, contemporary researchers have had to compile indexes from a range of approximate items (measures of trust, confidence in government, voting trends, social mobility,

and so on). Several excellent studies have identified useful measures of, and proxies for, social capital, however.

One measure is membership in informal and formal associations and networks. In developing countries generally, and in rural areas in particular, measures that capture the informal give-and-take through communitywide festivals, sporting events, and other traditional methods of fostering social connections are very important indicators of the underlying stocks of social capital. Based on data from a survey of 1,400 households in 87 villages across Tanzania (Narayan 1997), Narayan and Pritchett (1999) developed an index of social capital at the household and community levels that included density and characteristics of informal and formal groups and networks. The dimensions of this index included group functioning, financial and in-kind contributions to groups, participation in decisionmaking, and heterogeneity of membership. A series of measures was also constructed on interpersonal trust and changes over time. These measures demonstrated that social capital was indeed both social and capital, generating returns that exceeded those to human capital.

In tandem with the Tanzania study, studies of local institutions in three countries—Bolivia (Grootaert and Narayan 2000), Burkina Faso (Grootaert, Oh, and Swamy 1999), and Indonesia (Grootaert 1999)—looked at qualitative service delivery issues and quantified these variables. These studies demonstrated that the questionnaire items do in fact capture different dimensions of social capital at the household and community levels, that certain dimensions of social capital contribute significantly to household welfare, and that social capital is the capital of the poor The most important variables in these studies are density of associations, heterogeneity of membership in associations, and degrees of active participation in them

Another manifestation of social capital includes norms and values that facilitate exchanges, lower transaction costs, reduce the cost of information, permit trade in the absence of contracts, and encourage responsible citizenship and the collective management of resources (Fukuyama 1995). Inglehart's (1997) work on the World Values Survey is the most comprehensive effort in this area. The questions economists working on social capital find valuable are those on trust ("Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"). Knack and Keefer (1997), for example, use these data to show the positive relationship between trust and levels of investment in a country

Although research attempting to identify the nature of the relationships between social variables and development has recently proliferated, the quality of the data is less than ideal. With mounting pressure to provide simple measures of inherently complex and interdependent relationships, there is a danger that expectations will exceed capacity and that hastily assembled, poorly conceived measures will jeopardize the agenda they purport to serve. One way to strike the balance between quality and quantity measures is to unbundle social capital into its dimensions and to generate new data sets that are comparable across many countries.[2]

Four recent studies attempt to develop indices of social capital at the national or subnational levels. In the United States several new surveys of civic engagement are being conducted in addition to the work already collected in surveys of consumer preferences and changes in lifestyles. The National Commission on Philanthropy and Civic Renewal (1998), for example, has developed a National Index of Civic Engagement based on a sample of 1,000 respondents. This index includes five dimensions: the giving climate, community engagement, charitable involvement, the spirit of voluntarism, and active citizenship. Robert Putnam's Saguaro Seminar will soon launch the Social Capital Community Benchmark, a comprehensive survey of social capital in the United States (Putnam 2000).

In exploring the roots and determinants of Hindu and Muslim riots in India, Varshney (2000) focuses on the role of intercommunal networks. In cities where Hindus and Muslims have little interaction, Varshney shows that latent communal conflict has few channels for peaceful resolution and periodically descends into violence; in cities where association memberships overlap and everyday interactions are frequent, conflict is anticipated and dissipated. This research was based on six Indian cities carefully arranged in three matched pairs that were similar in terms of Hindu-Muslim demographic composition but dissimilar in that one city experienced recurrent riots whereas the other city remained calm. Varshney's work shows that diversity can be a source of strength where social ties transcend different community boundaries.

To assess social capital at the community level, Onyx and Bullen (forthcoming) developed a questionnaire for the state of New South Wales, Australia, from which they isolated eight underlying factors that constituted an individual's social capital: participation in the local community, proaction in a social context, feelings of trust and safety, neighborhood connections, connections with family and friends, tolerance of diversity, value of life, and work connections. Looking only at an individual's social capital score, the authors could predict the community to which the person belonged, thus raising the prospects for this instrument being used for planning and monitoring community development activities.

Building on this work, researchers are working to develop social capital instruments that can be used as diagnostic tools at the community level and across countries. Because the forms of social capital are society-specific and change over time, the instruments must focus on a range of dimensions of social capital (Narayan and Cassidy 1999). Such instruments have recently been introduced in Ghana and Uganda (Narayan 1998) and by the World Bank's Social Capital Initiative in Panama and India (Krishna and Shrader 1999).[3] Analyses of the data reveal that the dimensions underlying social capital are strikingly similar even when the context is quite different. The Ghana study draws on a sample of 1,471 rural and urban households, while the Uganda study focuses on 950 households in slum communities in Kampala. Factor analyses reveal a similar underlying structure and clustering of variables.

Implications for Development Theory and Policy

The concept of social capital offers a way to bridge sociological and economic perspectives and to provide potentially richer and better explanations of economic development. One important way it does this is by showing that the nature and exten of social interactions between communities and institutions shape economic perfor mance. This, in turn, has important implications for development policy, which ha long focused exclusively on an economic dimension. Similarly, understanding hov outside agencies can work to alleviate poverty in diverse and poorly understood communities remains one of the great challenges of development. A social capital perspective stresses that technical and financial soundness is a necessary but insufficien condition for acceptance of a project by poor communities.

Six broad recommendations can be offered for incorporating the concept of socia capital into development policy. First, for development interventions in all sector and at all levels (especially the country level), social institutional analysis should b used to identify correctly the range of stakeholders and their interrelations. Understanding how proposed policy interventions will affect the power and political interests of the stakeholders is a vital consideration, since all policy interventions occur i a social context characterized by a delicate mix of informal organizations, networks and institutions. The design of an intervention needs to pay special attention to th potential for dominant groups to mobilize in ways that undermine the public good

Second, it is critical to invest in the organizational capacity of the poor and to hel build bridges between communities and social groups. The latter is particularly important because many decisions affecting the poor are not made at the local level. T this end, the use of participatory processes can facilitate consensus-building and social interaction among stakeholders with diverse interests and resources. Findin ways and means by which to transcend social divides and build social cohesion an trust is crucial for economic development. One of the great virtues of the idea an discourse on social capital is that it provides a common language for these differen stakeholders, enabling them to communicate more easily with one another.

Third, a social capital perspective adds its voice to those calling for informatio disclosure policies at all levels to encourage informed citizenship and accountabilit of both private and public actors who purport to serve the public good. Fourth improvements in physical access and modern communications technology that ca foster information exchange across social groups should be emphasized to complement social interaction based on face-to-face interchange. Fifth, development interventions should be viewed through a social capital lens, and assessments of their impact should include the potential effects of the intervention on the social capital o poor communities. To reiterate, the social networks of the poor are one of the primar resources they have for managing risk and vulnerability, and outside agents therefor need to find ways to complement these resources, rather than substitute for them.

Finally, social capital should be seen as a component of orthodox development projects, from dams and irrigation systems to local schools and health clinics. Where poor communities have direct input into the design, implementation, management, and evaluation of projects, returns on investments and the sustainability of the project are enhanced (Esman and Uphoff 1984).

Conclusion

Although it is too soon to announce the arrival of a new development paradigm, it is not unreasonable to claim that a consensus is emerging about the importance of social relations in development. In unpacking the literature on social capital and development, a recurring message is that social relations provide opportunities for mobilizing other growth-enhancing resources, that social capital does not exist in a political vacuum, and that the nature and extent of the interactions between communities and institutions hold the key to understanding the prospects for development in a given society. The evidence supports the argument that social capital can be used to promote or to undermine the public good. This consideration suggests that one of the most important examples of social capital at work in the absence of formal insurance mechanisms and financial instruments is the use by the poor of social connections to protect themselves against risk and vulnerability.

In many respects the research on social capital is still in its early stages, but practitioners and policymakers cannot wait for researchers to know all there is to know before acting. Instead, all those involved should adopt a stance of learning by doing. This implies more rigorous evaluations of project and policy impact on social capital, more work on unbundling the mechanisms through which social capital works, and understanding the determinants of social capital itself. It also implies that practical lessons emerging from development projects can themselves be used to inform social capital theory.

It would be the ultimate irony if those people most interested in studying social capital and promoting its use in formulating development policy did not themselves foster trust, openness, and a willingness to share information, ideas, and opportunities in this field. Readers are invited to access, use, and contribute to the ongoing research on social capital.[4] It is only through collaborative efforts—with all that this entails regarding struggle, perseverance, negotiation, and mutual willingness to learn—that genuine progress will be made.

Notes

Michael Woolcock is a social scientist with the World Bank's Development Research Group and an adjunct lecturer in public policy at Harvard University Deepa Narayan is a lead social development

specialist in the Poverty Reduction and Economic Management Network at the World Bank. For helpful comments on an earlier draft of this paper, the authors thank John Blaxall, Jonathan Fox, Christiaan Grootaert, Bill Mulford, Vijayendra Rao, Anders Rudkvist, and anonymous reviewers

1. For citations on the first eight fields, see Woolcock (1998) and Foley and Edwards (1999). See also the database of articles on the World Bank's social capital Website, at <http://www.worldbank.org/poverty/scapital/library/index.htm#db>.

2. A number of recent survey instruments are available to researchers doing work in this field. See the World Bank's social capital Website, <http://www.worldbank.org/poverty/scapital/library/surveys.htm>

3 The World Bank's Social Capital Initiative is a $1.2 million dollar project sponsored by the government of Denmark. Several monographs produced for the initiative have been cited in this paper; these and several others can be downloaded at <http://www.worldbank.org/poverty/scapital/wkrppr/wrkppr htm>. These papers are currently being edited and prepared for formal publication

4 The World Bank's Social Capital Thematic Group Website contains instructions on how to receive our newsletter and join the e-mail discussion group. Go to <http://www worldbank.org/poverty/scapital/>.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems

Astone, Nan Marie, Constance Nathanson, Robert Schoen, and Young Kim. 1999. "Family Demography, Social Theory, and Investment in Social Capital." *Population and Development Review* 25(1):1–31

Barr, Abigail 1998 "Enterprise Performance and the Functional Diversity of Social Capital." Working Paper Series 98-1. University of Oxford, Institute of Economics and Statistics, Oxford, U K Processed

Berry, Sara. 1993. *No Condition Is Permanent. The Social Dynamics of Agrarian Change in Sub-Saharan Africa.* Madison, Wis.. University of Wisconsin Press.

Briggs, Xavier de Souza. 1998. "Brown Kids in White Suburbs: Housing Mobility and the Multiple Faces of Social Capital." *Housing Policy Debate* 9(1):177–221.

Brown, L David. 1998 "Creating Social Capital: Nongovernmental Development Organizations and Intersectoral Problem Solving." In Walter W. Powell and Elisabeth Clemens, eds., *Private Action and the Public Good* New Haven, Conn.· Yale University Press.

Brown, L David, and Darcy Ashman. 1996. "Participation, Social Capital, and Intersectoral Problem Solving African and Asian Cases." *World Development* 24(6):1477–95.

Burt, Ronald. 1992 *Structural Holes. The Social Structure of Competition.* Cambridge, Mass.· Harvard University Press.

———. 1997. "The Contingent Value of Social Capital." *Administrative Science Quarterly* 42 339–65.

——— 1998 "The Gender of Social Capital." *Rationality and Society* 10(1) 5-46.

Coleman, James. 1987 "Norms as Social Capital " In Gerard Radnitzky and Peter Bernholz, eds, *Economic Imperialism The Economic Method Applied outside the Field of Economics* New York Paragon House Publishers.

———. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94(Supplement):S95–S120.

———. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.

Collier, Paul. 1998. "Social Capital and Poverty." Social Capital Initiative Working Paper 4. World Bank, Social Development Department, Washington, D.C Processed

———. 1999. "The Political Economy of Ethnicity." In Boris Pleskovic and Joseph E. Stiglitz, eds , *Annual World Bank Conference on Development Economics 1998* Washington, D C . World Bank.

Collier, Paul, and Jan Willem Gunning 1999 "Explaining African Economic Performance." *Journal of Economic Literature* 37(March):64–111.

Dordick, Gwendolyn. 1997. *Something Left to Lose Personal Relations and Survival among New York's Homeless.* Philadelphia, Pa.· Temple University Press.

Easterly, William. 2000. "Happy Societies. The Middle Class Consensus and Economic Development." World Bank, Development Research Group, Washington, D C Processed

Escobar, Arturo 1995 *Encountering Development: The Making and Unmaking of the Third World* Princeton, N.J.: Princeton University Press

Esman, Milton, and Norman Uphoff 1984 *Local Organizations. Intermediaries in Rural Development.* Ithaca, N.Y.: Cornell University Press

Evans, Peter. 1992. "The State as Problem and Solution Predation, Embedded Autonomy, and Structural Change." In Stephan Haggard and Robert Kaufman, eds , *The Politics of Economic Adjustment· International Constraints, Distributive Conflicts, and the State* Princeton, N.J. Princeton University Press.

——— 1995 *Embedded Autonomy. States and Industrial Transformation* Princeton, N J.: Princeton University Press

———. 1996. "Government Action, Social Capital and Development Reviewing the Evidence on Synergy." *World Development* 24(6)·1119–32.

Fafchamps, Marcel, and Bart Minten. 1999 "Social Capital and the Firm: Evidence from Agricultural Trade." Social Capital Initiative Working Paper 17 World Bank, Social Development Department, Washington, D.C. Processed.

Foley, Michael, and Bob Edwards 1999. "Is It Time to Disinvest in Social Capital?" *Journal of Public Policy* 19 141–73.

Fox, Jonathan 1992. "Democratic Rural Development: Leadership Accountability in Regional Peasant Organizations " *Development and Change* 23(2):1–36.

Fox, Jonathan, and L. David Brown, eds. 1998 *The Struggle for Accountability· The World Bank, NGOs, and Grassroots Movements.* Cambridge, Mass. MIT Press.

Fukuyama, Francis. 1995 *Trust The Social Virtues and the Creation of Prosperity* New York· Free Press

Gittell, Ross, and Avis Vidal 1998 *Community Organizing Building Social Capital as a Development Strategy* Newbury Park, Calif. Sage Publications.

Granovetter, Mark. 1973 "The Strength of Weak Ties " *American Journal of Sociology* 78 1360–80.

——— 1995. "The Economic Sociology of Firms and Entrepreneurs " In Alejandro Portes, ed., *The Economic Sociology of Immigration. Essays on Networks, Ethnicity, and Entrepreneurship* New York. Russell Sage Foundation.

Grootaert, Christiaan 1999 "Social Capital, Household Welfare, and Poverty in Indonesia." Local Level Institutions Working Paper 6 World Bank, Social Development Department, Washington, D.C. Processed.

Grootaert, Christiaan, and Deepa Narayan. 2000 "Local Institutions, Poverty, and Household Welfare in Bolivia." Working Paper 9. World Bank, Social Development Department, Washington, D.C Processed.

Grootaert, Christiaan, Gi-Taik Oh, and Anand Swamy. 1999. "Social Capital and Development Outcomes in Burkina Faso." Working Paper 7. World Bank, Social Development Department, Washington, D.C. Processed.

Hall, Robert, and Charles Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114(1):83–116.

Hanifan, Lyda J. 1916. "The Rural School Community Center." *Annals of the American Academy of Political and Social Science* 67·130–38

Holzmann, Robert, and Steen Jorgensen. 1999. "Social Protection as Social Risk Management." Social Protection Discussion Paper 9904. World Bank, Human Development Network, Social Protection Team, Washington, D.C. Processed.

Homans, George. 1961. *Social Behavior. Its Elementary Forms.* New York: Harcourt, Brace and World

Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies* Princeton, N J : Princeton University Press

Isham, Jonathan. 1999. "The Effect of Social Capital on Technology Adoption: Evidence from Rural Tanzania." Paper presented at the annual meeting of the American Economic Association, New York. Processed.

Isham, Jonathan, and Daniel Kaufmann. 1999 "The Forgotten Rationale for Policy Reform The Productivity of Investment Projects." *Quarterly Journal of Economics* 114(1):149–84.

Isham, Jonathan, Deepa Narayan, and Lant Pritchett 1995 "Does Participation Improve Performance? Establishing Causality with Subjective Data " *The World Bank Economic Review* 9(2) 175–200

Jacobs, Jane. 1961. *The Life and Death of Great American Cities* New York: Random House.

Keck, Margaret, and Kathryn Sikkink 1998. *Activists beyond Borders Advocacy Networks in International Politics* Ithaca, N.Y . Cornell University Press

Knack, Stephen 1999. "Social Capital, Growth, and Poverty A Survey of Cross-Country Evidence" Social Capital Initiative Working Paper 7. World Bank, Social Development Department, Washington, D C Processed.

Knack, Stephen, and Philip Keefer 1995 "Institutions and Economic Performance Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics* 7(3)·207–27

———. 1997 "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation" *Quarterly Journal of Economics* 112.1251–88

Kozel, Valerie, and Barbara Parker 2000 "Integrated Approaches to Poverty Assessment in India" In Michael Bamberger, ed , *Integrating Quantitative and Qualitative Research in Development Projects* Washington, D C · World Bank.

Krishna, Anirudh, and Elizabeth Shrader. 1999 "Social Capital Assessment Tool " Social Capital Initiative Working Paper 22 World Bank, Social Development Department, Washington, D C Processed

Loury, Glenn C 1977 "A Dynamic Theory of Racial Income Differences." In P. A. Wallace and A LeMund, eds , *Women, Minorities, and Employment Discrimination* Lexington, Mass. Lexington Books

Massey, Douglas. 1998. "March of Folly. U.S. Immigration Policy after NAFTA." *The American Prospect* 37.22–33.

Massey, Douglas, and Karin Espinosa 1997 "What's Driving Mexico-U S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102(4).939–99

Moore, Mick. 1997. "Societies, Polities, and Capitalists in Developing Countries. A Literature Survey." *Journal of Development Studies* 33(3)·287–363

Moser, Caroline. 1996. *Confronting Crisis: A Comparative Study of Household Responses to Poverty and Vulnerability in Four Poor Urban Communities* Environmentally Sustainable Development Studies and Monographs Series 8. Washington, D C.· World Bank

Narayan, Deepa. 1995. "Designing Community-Based Development." Social Development Paper 7. World Bank, Environmentally and Socially Sustainable Development Network, Washington, D.C Processed.

———. 1997. "Voices of the Poor: Poverty and Social Capital in Tanzania." Environmentally Sustainable Development Monograph 20. World Bank, Washington, D C.

———. 1998. "Social Capital Survey in Ghana—Preliminary Results." World Bank, Poverty Reduction and Economic Management Network, Washington, D C Processed.

———. 1999. "Bonds and Bridges Social Capital and Poverty " Policy Research Working Paper 2167. World Bank, Poverty Reduction and Economic Management Network, Washington, D C. Processed.

Narayan, Deepa, and Michael Cassidy 1999 "A Dimensional Approach to Measuring Social Capital: Development and Validation of a Social Capital Inventory " World Bank, Poverty Reduction and Economic Management Network, Washington, D.C. Processed.

Narayan, Deepa, and David Nyamwaya 1996 "Learning from the Poor A Participatory Poverty Assessment in Kenya " Environment Department Papers, Participation Series 34 World Bank, Social Policy and Resettlement Division, Washington, D C Processed.

Narayan, Deepa, and Lant Pritchett. 1999. "Cents and Sociability· Household Income and Social Capital in Rural Tanzania " *Economic Development and Cultural Change* 47(4) 871–97

Narayan, Deepa, and Talat Shah 1999 "Gender Inequity, Poverty, and Social Capital " Background paper prepared for the Policy Research Report on Gender and Development World Bank, Development Research Group, Washington, D.C Processed

Narayan, Deepa, with Raj Patel, Kai Schafft, Anne Rademacher, and Sarah Koch-Schulte 2000 *Voices of the Poor Can Anyone Hear Us? Voices from 46 Countries* Washington, D C . World Bank

National Commission on Philanthropy and Civic Renewal 1998. "National Index of Civic Engagement " University of Connecticut, Storrs, Conn

North, Douglass C 1990 *Institutions, Institutional Change, and Economic Performance* New York. Cambridge University Press

Onyx, Jenny, and Paul Bullen Forthcoming "Measuring Social Capital in Five Communities " *Journal of Applied Behavioral Science*

Platteau, Jean-Philippe 1994 "Behind the Market Stage Where Real Societies Exist" (Parts I and II). *Journal of Development Studies* 30.533–77 and 753–817

Portes, Alejandro, ed 1995 *The Economic Sociology of Immigration Essays on Networks, Ethnicity, and Entrepreneurship* New York: Russell Sage Foundation

——— 1997. "Neoliberalism and the Sociology of Development " *Population and Development Review* 23(2)·229–59

——— 1998 "Social Capital Its Origins and Applications in Contemporary Sociology." *Annual Review of Sociology* 24:1–24.

Portes, Alejandro, and Patricia Landolt. 1996 "The Downside of Social Capital " *The American Prospect* 26 (May-June).18–21, 94

Portes, Alejandro, and Julia Sensenbrenner 1993 "Embeddedness and Immigration. Notes on the Social Determinants of Economic Action " *American Journal of Sociology* 98(6) 1320–50

Putnam, Robert 1993 *Making Democracy Work Civic Traditions in Modern Italy* Princeton, N J Princeton University Press.

———. 1995. "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America." *PS: Political Science and Politics* (December):664–83.

———. 2000. *Bowling Alone: The Collapse and Revival of American Community* New York. Simon and Schuster.

Rodrik, Dani. 1998. "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses." NBER Working Paper 6350. National Bureau of Economic Research, Cambridge, Mass

———. 1999. *Making Openness Work.* Baltimore, Md.: Johns Hopkins University Press.

Rose, Richard. 1998. "Getting Things Done in an Anti-Modern Society: Social Capital Networks in Russia." Social Capital Initiative Working Paper 8. World Bank, Social Development Department, Washington, D.C. Processed.

Rubio, Mauricio. 1997. "Perverse Social Capital: Some Evidence from Colombia." *Journal of Economic Issues* 31(3):805–16.

Schafft, Kai. 1998. "Grassroots Development and the Reconfiguration of Local Political Institutions Local Minority Self-Governance as a Political and Economic Resource for Hungary's Roma Population." Paper presented at the Graduate Student Conference on International Affairs, George Washington University, Washington, D.C Processed.

Seeley, John R., Alexander R. Sim, and Elizabeth W. Loosley. 1956. *Crestwood Heights A Study of the Culture of Suburban Life* New York: Basic Books.

Skocpol, Theda. 1995. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States.* Cambridge, Mass.: Harvard University Press.

——— 1996. "Unraveling from Above." *The American Prospect* 25:20–25.

Temple, Jonathan. 1998. "Initial Conditions, Social Capital, and Growth in Africa." *Journal of African Economies* 7(3):309–47.

Temple, Jonathan, and Paul Johnson. 1998 "Social Capability and Economic Growth." *Quarterly Journal of Economics* 113(3):965–90.

Tendler, Judith. 1997. *Good Government in the Tropics.* Baltimore, Md.: Johns Hopkins University Press.

United Nations. 1951 *Measures for the Economic Development of Underdeveloped Countries* New York: United Nations, Department of Social and Economic Affairs

Uphoff, Norman. 1992. *Learning from Gal Oya: Possibilities for Participatory Development and Post-Newtonian Social Science* Ithaca, N Y.: Cornell University Press

van Bastalaer, Thierry. 1999. "Does Social Capital Facilitate the Poor's Access to Credit? A Review of the Microeconomic Literature." Social Capital Initiative Working Paper 8. World Bank, Social Development Department, Washington, D C Processed

Van Dijk, Meine Pieter, and Roberta Rabellotti. 1997. *Enterprise Clusters and Networks in Developing Countries* London: Frank Cass.

Varshney, Ashutosh. 2000. *Ethnic Conflict and Civic Life Hindus and Muslims in India.* New Haven, Conn.. Yale University Press.

White, T Anderson, and Glenn Smucker 1998. "Social Capital and Governance in Haiti· Traditions and Trends " In *The Challenges of Poverty Reduction* Report 17242-HA Washington, D C World Bank.

Wilson, William Julius. 1987. *The Truly Disadvantaged* Chicago, Ill . University of Chicago Press

———. 1996. *When Work Disappears The World of the New Urban Poor.* New York: Knopf

Woolcock, Michael. 1998. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework." *Theory and Society* 27(2) 151–208.

———. 1999. "Learning from Failures in Microfinance: What Unsuccessful Cases Tell Us about How Group-Based Programs Work." *American Journal of Economics and Sociology* 58(1):17–42

———. 2000. "Managing Risk, Shocks, and Opportunities in Developing Economies: The Role of Social Capital." In Gustav Ranis, ed., *Dimensions of Development* New Haven, Conn.: Yale Center for International and Area Studies.

———. Forthcoming. *Using Social Capital: Getting the Social Relations Right in the Theory and Practice of Economic Development* Princeton, N.J. Princeton University Press.

World Bank. 1989. "Staff Appraisal Report, Rwanda, Agricultural Services Project." Report 7599-RW Agricultural Operations Division, South-Central and Indian Ocean Department, Africa Region. Processed

———. 1998. *Assessing Aid What Works, What Doesn't, and Why* New York Oxford University Press.

World Development. 1996. Special issue on social capital 24(6).

What Should the World Bank Think about the Washington Consensus?

John Williamson

The phrase "Washıngton Consensus" has become a famılıar term ın development polıcy cırcles in recent years, but it ıs now used in several different senses, causıng a great deal of confusion. In thıs artıcle the author dıstinguishes between his original meanıng as a summary of the lowest common denomınator of polıcy advıce addressed by the Washington-based ınstitutıons (ıncluding the World Bank) and subsequent use of the term to signıfy neoliberal or market-fundamentalıst polıcıes. He argues that the latter polıcıes could not be expected to provıde an effectıve framework for combatıng poverty but that the original advıce is still broadly valid. The article dıscusses alternative ways of addressing the confuıon It argues that any policy manifesto designed to eliminate poverty needs to go beyond he orıgınal version but concludes by cautionıng that no consensus on a wıder agenda urrently exists.

'en years ago I ınvented the term "Washıngton Consensus" to refer to the lowest ommon denominator of polıcy advice beıng addressed by the Washington-based ıstıtutions to Latin American countries as of 1989 (Williamson 1990). While it is ɔlly to become famous for coinıng a term that reverberates around the world, I have ɔng been doubtful about whether my phrase served to advance the cause of rational conomic policymaking. My inıtıal concern was that the phrase invıted the interpreatıon that the liberalizing economic reforms of the past two decades were imposed y Washington-based institutions (for example, see Stewart 1997) rather than havıg resulted from the process of ıntellectual convergence that I believe underlies the eforms.[1] Richard Feinberg's "universal convergence" (ın Wıllıamson 1990) or Jean Vaelbroeck's "one-world consensus" (Waelbroeck 1998) would have been a much etter term for the intellectual convergence that I had in mind.

I have gradually developed a second and more sıgnificant concern, however. I find ıat the term has been invested with a meanıng that is significantly different from ıat which I had intended and is now used as a synonym for what is often called

he World Bank Research Observer, vol. 15, no 2 (August 2000), pp 251–64

"neoliberalism" in Latin America, or what Geeorge Soros (1998) has called "market fundamentalism." When I first came across this usage, I asserted that it was a misuse of my intended meaning. I had naïvely imagined that just because I had invented the expression, I had some sort of intellectual property rights that entitled me to dictate its meaning, but in fact the concept had become public property.

The battle of economic ideas, as McCloskey (1998) has argued, is fought to a significant extent with rhetoric. The use of a term with dual meanings and strong ideological overtones can therefore pose serious dangers not only of misunderstanding but also of inadvertently prejudicing policy objectives. Specifically, there is a real danger that many of the economic reforms favored by international development institutions—notably macroeconomic discipline, trade openness, and market-friendly microeconomic policies—will be discredited in the eyes of many observers, simply because these institutions are inevitably implicated in views that command a consensus in Washington and the term "Washington Consensus" has come to be used to describe an extreme and dogmatic commitment to the belief that markets can handle everything.

The objective of this article is to consider what should be done to minimize the damage to the cause of intellectual understanding, and therefore of rational economic reform, that is being wrought by the current widespread use of the term "Washington Consensus" in a sense different from that originally intended. Would it be productive, for example, to insist that the original usage is the correct one? Or should one simply refuse to debate in these terms? Is it possible to escape by declaring fidelity to some "post–Washington Consensus"? The first stage in answering these questions is a careful examination of the semantic issues involved

The Original Version

My original paper (Williamson 1990) argued that the set of policy reforms that most of official Washington thought would be good for Latin American countries could be summarized in 10 propositions:

- Fiscal discipline
- A redirection of public expenditure priorities toward fields offering both high economic returns and the potential to improve income distribution, such as primary health care, primary education, and infrastructure
- Tax reform (to lower marginal rates and broaden the tax base)
- Interest rate liberalization
- A competitive exchange rate
- Trade liberalization
- Liberalization of inflows of foreign direct investment

- Privatization
- Deregulation (to abolish barriers to entry and exit)
- Secure property rights.

The need for the first three reforms is, so far as I am aware, widely accepted among economists. Nevertheless, when I reviewed the progress that Latin American countries had made in implementing the recommended set of policies several years later (Williamson 1996), it appeared that the least progress had come in redirecting public expenditure priorities. The other seven reforms have stimulated a measure of controversy and therefore merit comment.

In my original paper I specified interest rate liberalization as the fourth reform. I am now well aware that many economists have reservations about that formulation. As a matter of fact, I have such reservations myself: in Williamson and Mahar (1998) interest rate liberalization is identified as merely one of six dimensions of financial liberalization. Moreover, Stiglitz (1994) has argued that interest rate liberalization should come toward the end of the process of financial liberalization, inasmuch as a ceiling on the deposit interest rate (equal to the Treasury bill rate, he suggests) might provide a constraint on gambling for redemption. I find this argument persuasive and long ago changed my description of the fourth element of the Washington Consensus to financial liberalization. More recently Stiglitz (1998) has expressed a much more basic objection to financial liberalization, arguing that the success of some East Asian countries stemmed importantly from their policy of directing credit to particular industries rather than allowing the market to determine the allocation of credit. That argument is highly contentious, especially in the aftermath of the East Asian economic crisis of 1997–98.

My fifth choice—a competitive exchange rate—was not, I have concluded, an accurate report of Washington opinion. I suspect that by 1989 a majority of economists, in Washington as elsewhere, were already in favor of either firmly fixed or freely floating exchange rates and hostile to the sort of intermediate regime that in my judgment gives the best promise of maintaining a competitive exchange rate in the medium term. (My own preference remains an intermediate regime of limited flexibility, provided that excludes an old-fashioned adjustable peg, even if such a regime is more likely to spawn speculative pressures than a floating rate.) But note that the East Asian countries did by and large achieve and maintain competitive exchange rates, at least before about 1996 (and even after 1996 only Thailand failed to do so).[2]

My sixth reform was trade liberalization. Here I see little reason to doubt that I reported accurately on opinions in the international financial institutions and the central economic agencies of the U.S. government (although parts of Congress and the Department of Commerce are not noted for their dedication to liberal trade). But this is another area where critics can rightly claim that the policies that nurtured

the East Asian miracle were, at least in some countries, at odds with the policies endorsed in the Washington Consensus. Much the same is true of foreign direct investment, except that the East Asian economies were less hostile to a policy of openness; only the Republic of Korea rejected most foreign direct investment during the years of the miracle.

Privatization commanded a lot of support in Washington, where it had been put on the international agenda by James Baker when he was secretary of the U.S. Treasury, in his speech to the World Bank–International Monetary Fund Annual Meetings in Seoul in 1985. Privatization was controversial in much of the rest of the world, where one's attitude to public versus private ownership had long been the litmus test for qualifying as left-wing or right-wing. Deregulation was rather less politically polarizing: it had been initiated by the centrist Carter administration in the United States, rather than by the right-wing Thatcher government that pioneered privatization in the United Kingdom. Deregulation, however, was not a policy that reverberated in East Asia, where the industrial policies pursued in some (though not all) countries ran very much in the opposite direction. The notion of the importance of secure property rights had come both from Chicago's law and economics school and from the work of Hernando de Soto in Peru. The concept was presumably offensive to those who resisted the advance of the market economy, but this breed was extinct in Washington by 1989 (if, indeed, it had ever existed there). My impression is that the institution of private property was somewhat more securely entrenched in East Asia than in most of the rest of the developing world.

So much for the content of my version of the Washington Consensus. What inspired it? In an immediate sense, it originated from an attempt to answer a question posed to me by Hans Singer during a seminar at the Institute for Development Studies: what were these "sensible" policies that were being pursued in Latin America (and that I was arguing justified approval of the Brady Plan to provide these countries with debt relief)? In a more profound sense, my effort was an attempt to distill which of the policy initiatives that had emanated from Washington during the years of conservative ideology had won inclusion in the intellectual mainstream rather than being cast aside once Ronald Reagan was no longer on the political scene.[3] Taking an even longer perspective, my version of the Washington Consensus can be seen as an attempt to summarize the policies that were widely viewed as supportive of development at the end of the two decades when economists had become convinced that the key to rapid economic development lay not in a country's natural resources or even in its physical or human capital but, rather, in the set of economic policies that it pursued.

Let me emphasize that the Washington Consensus as I conceived it was in principle geographically and historically specific, a lowest common denominator of the reforms that I judged "Washington" could agree were needed in Latin America as of 1989. But in practice there would probably not have been a lot of difference if I had

undertaken a similar exercise for Africa or Asia, and that still seemed to be the case when I revisited the topic (with regard to Latin America) in 1996 (Williamson 1997). This doubtless made it easier for some to interpret the Washington Consensus as a policy manifesto that its adherents supposedly believed to be valid for all places and at all times.

Current Usage

The following is a selection of recent definitions of the Washington Consensus that I happened to stumble across. (I have undertaken no bibliographic research to compile this list.)

"A die-hard liberalization advocate (or a Washington-consensus believer). . . . " (Ito 1999)

". . . the self-confident advice of the 'Washington consensus'—free-up trade, practice sound money, and go home early. . . ." (Vines 1999)

" . . . the Washington Consensus: policy prescriptions based on free market principles and monetary discipline." (Hamada 1998)

"The Washington Consensus had the following message: 'Liberalize as much as you can, privatize as fast as you can, and be tough in monetary and fiscal matters.'" (Kolodko 1998)

"The bashing of the state that characterized the policy thrust of the Washington Consensus. . . ." (United Nations 1998)

"This new imperialism, codified in the 'Washington Consensus'. . . ." (Alam 1999)

"The Brazilian crisis has reignited the debate over the so-called Washington Consensus on the creation of a laissez-faire global economy." (Rajan 1999)

In none of these examples is my phrase used in the sense that I originally intended. On the contrary, when I coined the term in 1989, the market fundamentalism of Reagan's first term had already been superseded by the return of rational economic policymaking, and one could discern which ideas were going to survive and which were not (monetary discipline but not monetarism; tax reform but not tax-slashing; trade liberalization but maybe not complete freedom of capital movements; deregulation of entry and exit barriers but not the suppression of regulations designed to protect the environment).

How is it that a term intended to describe a technocratic policy agenda that survived the demise of Reaganomics came to be used to describe an ideology embracing the most extreme version of Reaganomics? The closest I can come to understanding this is to note that my version of the Washington Consensus did indeed focus principally on policy reforms that reduced the role of government, such as privatization and the liberalization of trade, finance, foreign direct investment, and entry and exit. It did this because the orthodoxy of the generation whose ideas were embodied in

the practices being challenged in 1989 had been much more statist than was by then regarded as advisable, and hence the policy reforms that were needed at that time were all in the direction of liberalization. This need for liberalization did not necessarily imply a swing to the opposite extreme of market fundamentalism and a minimalist role for government, but such boring possibilities were repressed in the ideological debates of the 1990s. For it is certainly true that the Washington Consensus came to be used to describe an ideological position, a development that Naim (2000) argues resulted from the world's acute need for a new ideology to provide a focus for debate in place of the god that had failed. My qualifications about the Washington Consensus being an agenda for a specific part of the world at a particular moment of history were quickly forgotten, as the search for a new ideology, to endorse or to hate, was perceived to have succeeded. Ravi Kanbur argues that the staffs of the Bretton Woods institutions perceived themselves as storming the citadels of statism, which led them as a negotiating ploy to demand more in the way of liberalizing reforms than they really expected to achieve—a tactic that led citizens in the World Bank's client countries to identify these institutions with something closer to market fundamentalism than the institutions really believed in.

The term's use as a synonym for market fundamentalism appears to be the dominant, but not the only, current usage. Many Bank staff members, including those who wrote *Beyond the Washington Consensus: Institutions Matter* (Burki and Perry 1998), still use the term in the way that I intended, and I think most of them would endorse the reform agenda to which I had applied the term as a reasonably accurate and appropriate summary of what the Bank and other agencies concerned with the promotion of development were, and should have been, advising countries to do

Joseph Stiglitz, formerly the World Bank's chief economist, recently used the term in the alternative, neoliberal, sense (1999b). This at least makes it clear that he was not attacking his colleagues when he spoke of reviewing "the major ways in which . . . the 'Washington Consensus' doctrines of transition, failed. . ." (Stiglitz 1999a:4) He proceeded to question the priority given to rapid privatization and the lack of attention to establishing competition or building social and organizational capital, and later he spoke of "the standard form of voucher privatization promoted by the Washington Consensus. . . ." I am not aware that Washington has ever displayed any particular preference for voucher privatization; certainly this was not a theme of the 1996 *World Development Report* (World Bank 1996), which dealt with the transition. I agree with Stiglitz on the substantive questions he raises: one can put too much emphasis on rapid privatization, and it is more important to do it right than to do it quickly; I agree that the great merit of privatization is that it can be used to further competition; I am skeptical about voucher privatization; and I think I agree about the importance of social and organizational capital, if I understand what the words mean. (I would describe them as social cohesion and good institutions, respectively.) What I do not understand is what is gained by describing these sensible

ideas as refuting a doctrine described by a term that many people in the Bank regard as providing a useful summary of the advice the Bank dispenses.

Do Washington Consensus Policies Promote Poverty Reduction?

The answer, quite obviously, depends on which interpretation of the Washington Consensus one is referring to. The popular, or populist, interpretation of the Washington Consensus, meaning market fundamentalism or neoliberalism, refers to laissez-faire Reaganomics—let's bash the state, the markets will resolve everything. I would not subscribe to the view that such policies offer an effective agenda for reducing poverty. We know that poverty reduction demands efforts to build the human capital of the poor, but the populist interpretation fails to address that issue. We know that an active policy to supervise financial institutions is needed if financial liberalization is not to lead to financial collapse, which invariably ends up using tax revenues to write off bank loans that were made to the relatively rich. And some measure of income redistribution would be recommended by any policy that was primarily directed at reducing poverty rather than simply maximizing growth, but market fundamentalists rule out all income redistribution as plunder.

A plausible alternative concept would be that the Washington Consensus consists of the set of policies endorsed by the principal economic institutions located in Washington: the U.S. Treasury, the Federal Reserve Board, the International Monetary Fund, and the World Bank I would argue that the policies these institutions advocated in the 1990s were inimical to the cause of poverty reduction in emerging markets in at least one respect: their advocacy of capital account liberalization. This was, in my view, the main cause of the contagion that caused the East Asian crisis to spread beyond Thailand and that resulted in a tragic interruption of the poverty reduction those countries had achieved (Williamson 1999). (I did not include full capital account liberalization in my version of the Washington Consensus because I did not believe it commanded a consensus, if only because I could not believe I was the only person in Washington who feared that capital account liberalization could precipitate a tragedy such as that which occurred in East Asia.)

My version of the Washington Consensus began with the proposition that the inflation caused by lack of fiscal discipline is bad for income distribution. The second reform specifically involved redirecting public expenditure toward primary health and education, that is, toward building the human capital of the poor. Tax reform can be distributionally neutral or even progressive. A competitive exchange rate is key to nurturing export-led and crisis-free growth and is hence in the general interest, including that of the poor. Trade liberalization, certainly in low-income, resource-poor countries, tends to be pro-poor because it increases the demand for unskilled

labor and decreases the subsidies directed to import-competing industries that use large volumes of capital and employ small numbers of workers, many of them highly skilled. Foreign direct investment helps raise growth and spread technology, provided that import protection is not excessive, so that the case of immiserizing growth does not arise (Brecher and Diaz-Alejandro 1977).[4] The impact of privatization depends very much on how it is done: the sort of insider-voucher privatization that occurred in Russia allows the plunder of state assets for the benefit of an elite, but a well-conducted privatization with competitive bidding can raise efficiency and improve the public finances with benefits to all, including the poor. Deregulation in general involves the dismantling of barriers that protect privileged elites (even if some of them, like trade unionists, have difficulty thinking of themselves as an elite), and hence there is a strong presumption that it will be pro-poor. Private property rights are certainly a defense primarily for those who have private property, but the improvement of such rights is nonetheless very likely to be pro-poor because these are the people who find themselves unable to defend their property when property rights are ill-defined (for example, Hernando de Soto's squatters on the periphery of Lima).

I have omitted one of the ten reforms from the preceding list: financial and interest rate liberalization. This is the primary focus of Stiglitz's criticisms when he refers to something that I can recognize as akin to my version of the Washington Consensus. I have realized for some time (see Williamson 1996) that my first formulation was flawed in that it neglected financial supervision, without which financial liberalization seems all too likely to lead to improper lending and eventually to a crisis that requires the taxpayers to pick up the losses from making bad loans (Williamson and Mahar 1998). But should economists therefore endorse the view that directed lending as pursued in some—though not all—East Asian countries is pro-growth and thus ultimately pro-poor? On this issue, at least, I would have thought that the East Asian crisis, especially in Korea, should have tempered economists' enthusiasm for the practice. The high debt-equity ratios that resulted from directed lending were certainly among the causes of the financial fragility that deepened the impact of the crisis.

Thus most of the reforms embodied in my version of the Washington Consensus are at least potentially pro-poor. In some cases this conclusion is sensitive to the way in which reform is implemented: that is certainly true of tax reform, privatization and, above all, financial liberalization. But I see no reason why the World Bank should back away from endorsing my version of the Washington Consensus in view of its reaffirmation of poverty reduction as its overarching mission. That is not to claim that the Washington Consensus, in any version, constituted a policy manifesto adequate for addressing poverty. My version quite consciously eschewed redistributive policies, taking the view that Washington had not reached a consensus on their

desirability. But time has moved on, and we are now looking to *World Development Report 2000/01* for an outline of the policies needed to supplement my version of the Washington Consensus in a world that takes poverty reduction seriously.

The Semantic Dilemma

One can react to the semantic dilemma posed by the different definitions currently in use in three possible ways. Consider these alternatives:

- *Insist on the original usage.* Insist that my version of the Washington Consensus is the only correct and legitimate interpretation, as a corollary of which the term will (with the qualifications noted above) be recognized as pro-poor. This alternative strikes me as both presumptuous and unrealistic: once a term has escaped into the public domain, one cannot dictate the reestablishment of a common usage. The likely result would be a perpetuation of the public confusion that I am attempting to address.
- *Abandon the term.* Refuse to debate in the terms that have been so compromised by the widespread adoption of the "populist" definition. I cannot imagine that this approach would end the populist use of the term; it would simply be a cop-out.
- *Endorse a post–Washington Consensus.* A more promising strategy has been adopted at least twice within the Bank. In 1998 the Latin America Regional Office of the World Bank issued a policy document that favored going beyond the Washington Consensus (Burki and Perry 1998). Stiglitz did almost the same, semantically at least, in urging a post–Washington Consensus in his lecture to the World Institute for Development Economics Research in January 1998.

When I first came across this approach, I thought it implied that the reforms included in the Washington Consensus were necessary but not sufficient for promoting development, an idea that seemed eminently reasonable. Clearly the Bank today would want to go further and endorse a wider array of antipoverty instruments than was able to command a consensus in 1989, when the most I thought I could legitimately include was the promotion of public expenditure on primary health and education.[5]

In their book, Burki and Perry (1998) explicitly refer to my version of the Washington Consensus and assert that the widespread implementation of the "first-generation" reforms it prescribed was paying off in Latin America in resumed growth and an end to high inflation. They noted that the reforms had not been equally effective in reducing poverty and inequality, which they argued demonstrated a

"need to focus on improving the quality of investments in human development, promoting the development of sound and efficient financial markets, enhancing the legal and regulatory environments (in particular, deregulating labor markets and improving regulations for private investment in infrastructure and social services), [and] improving the quality of the public sector (including the judiciary) . . ." (p. 4). This is an agenda dominated by institutional reform, which is indeed what has become known in Latin America as the second-generation reform agenda (Naim 1995).

It is not equally obvious why Stiglitz would want to propagate a post–Washington Consensus that implied endorsing and extending the original version, given his interpretation of what was included in it. In fact, the Stiglitz version of a post–Washington Consensus does not endorse any version of the original. He is advocating a policy package that is intended to supersede the Washington Consensus altogether. His new policy package is asserted to differ from the original in two dimensions.

First, he argues that the implicit policy objective underlying the Washington Consensus is inadequate. In addition to pursuing economic growth, the objectives should include "sustainable development, egalitarian development, and democratic development." In other words, he believes that policy objectives should include the state of the environment, income distribution, and democracy, as well as per capita gross national product. I find those objectives much more congenial than a single-minded preoccupation with economic growth, although I am not sure that the World Bank could formally endorse the pursuit of democracy (its Articles do, after all, forbid its involvement in politics).[6] Second, in addition to expanding the objectives, Stiglitz argues that it is necessary to pursue "sound financial regulation, competition policy and policies to facilitate the transfer of technology and transparency" to make markets work in a way that will support development.

I have a somewhat different view of what should be added to the Washington Consensus to make it a policy manifesto supportive of egalitarian, environmentally sensitive development. I agree that financial regulation (prudential supervision) is crucial and that transparency is a useful complement to supervision in achieving appropriate conduct of financial institutions. Moreover, competition is a natural complement to deregulation in promoting a well-functioning market economy (although a liberal import regime is the most effective competition policy in tradables, as Srinivasan argues in his comment in this volume). I would not have included technology transfer in such a manifesto, although I would have no objection to including institutional changes that seemed likely to promote technology transfer if I were reasonably confident that I knew what these changes were (besides accepting foreign direct investment). Similarly, I would consider it desirable to include policies focused on improved environmental conditions, although I am not sure that I would know how to select policy measures at a comparable level of generality to my 10

original points. But my emphasis would have been different; I would have focused much more generally on institutions. To explain why, let me offer a brief history of postwar development thinking.

In the first wave of theorizing about economic development, from the 1940s to the early 1960s, economists saw the accumulation of physical capital as the key to development (as reflected in the Harrod-Domar model, the Lewis model, and the two-gap model). The second phase recognized that human capital provided another and more inelastic constraint on development, a constraint that explained why Europe and Japan had recovered from World War II so rapidly, when growth in developing countries had been lagging despite the adoption of development policies and the beginning of large-scale aid. The third phase, which started about 1970 with the work of Little, Scitovsky, and Scott (1970) and Balassa (1970), emphasized that the policy environment influenced the level and dominated the productivity of investment. The Washington Consensus attempted to summarize the outcome of this debate on the policies that were conducive to economic development. The major advance of the 1990s stemmed from recognition that the central task of the transition from communist to market-based economies involved building the institutional infrastructure of a market economy. This realization was complemented by a growing recognition that bad institutions can sabotage good policies. This viewpoint was reflected in Stiglitz's (1999a) remarks on the transition, in Naim's (1995) work on supplementing the Washington Consensus, in Burki and Perry (1998), in the *World Development Reports* of 1997 and 1998, and in the World Bank's decision to launch a crusade against corruption.

What should one make of the idea of launching a post–Washington Consensus? I would not be happy at such a move if it were interpreted to imply a rejection of "the" Washington Consensus, although I would have no problem if it involved rejection of the populist, or market-fundamentalist, version. But it seems a somewhat odd crusade. The time of the original consensus, 1989, was an unusual period in that the ideological battles of the Reagan era, not to mention the cold war battle between capitalism and communism, were passing into history, leaving in their wake an unusually wide measure of agreement that several rather basic ideas of good economics were not only desirable but of key importance in the current policy agenda of at least one region—Latin America. Currently, there is no similar coalescing of views, certainly not on the wider agenda that Stiglitz has laid out. (Consensus on egalitarianism? With aid fatigue threatening the future of the International Development Association? On environmental sustainability? In a world where the U.S. Senate refuses even to consider ratifying the Kyoto Protocol?) I agree, rather, with Tim Geithner (1999:8): "I don't think anyone believes there is some universal model that can or should be imposed on the world—Washington consensus, post Washington consensus, or not."

Resolving the Dilemma

Let me conclude by laying out my own ideas on how to resolve the dilemma.

- There is little merit in attacking abstract, undefined concepts that are interpr to mean whatever the author momentarily decides they mean. It is better to out those concepts that are being criticized and debate policies on the bas their merits.
- The World Bank should recognize that the term Washington Consensus been used in very different ways. One summarizes policies that are pro-p another describes a policy stance that offers the poor very little and warrant support.
- It is appropriate to go beyond the Washington Consensus by emphasizing importance of the institutional dimension as well as of the sort of pol embodied in the original version of the Washington Consensus—policies will promote an equitable distribution of income as well as a rapid growt income.
- The hopeless quest to identify a consensus where there is none shoul abandoned in favor of a debate on the policy changes needed to achie rounded set of objectives encompassing at least the level, growth, distribution of income, as well as preservation of a decent environment.

The Bank will do the cause of economic development a great service if it frame future debate in these terms. Admittedly my suggestions do not answer pleas for a new ideology that would more adequately reflect the goals of the mult eral development banks and that might thus increase the chance of establishing l ownership of the sort of economic policy stance conducive to rapid and equit growth. Let me plead in defense that I am not a suitable person to launch an ic ogy, inasmuch as Naím (2000) characterizes an ideology as a thought-economi device and I actually believe that thinking is more desirable than economizing thought.

Notes

John Williamson is senior fellow at the Institute for International Economics This article was wr as a background paper for *World Development Report 2000/01.* The author is indebted to the pa pants in a session at which an early version of the paper was discussed, notably Ravi Kanbu Moisés Naím.

1. This intellectual convergence was the result of the collapse of communism, which resulte from machinations of the Bretton Woods institutions, or even of the U.S Central Intelligence Ag but because socialism does not work except in a simple economy, and even then it seems to worked reasonably well only when large numbers of people were inspired with revolutionary ze

2 Exchange rate policy is the one topic on which I have a serious difference of view with T. N Srinivasan's comment that accompanies this paper The term "competitive exchange rate" originated with Bela Balassa and signifies a rate that is either at, or undervalued relative to, its long-run equilibrium. I do not regard measuring the latter as an exercise in futility; see Hinkle and Montiel (1999) for evidence that other people in the Bank do not either. I dissent from the consensus Srinivasan proclaims that holds that only currency boards and freely floating rates offer viable regimes. For further details, see Williamson (forthcoming)

3. In trying to identify policies from the Reagan-Thatcher era that had not won consensus support, I wrote in 1996: "it [the Washington Consensus] did not declare that the only legitimate way to restore fiscal discipline was to slash government expenditure; it did not identify fiscal discipline with a balanced budget, it did not call for overall tax cuts; it did not treat as plunder the taxes raised to redistribute income; it did not say that exchange rates had to be either firmly fixed or freely floating; it did not call for the proscription of capital controls, it did not advocate competitive moneys or argue that the money supply should grow at a fixed rate" (Williamson 1997:50)

4. Growth of output of a heavily protected product can immiserize a country if the resources used in production exceed the social value of the output

5. However, in commenting on my paper, Stanley Fischer (then the Bank's chief economist) argued that I could and should have gone further: "Emphasis on poverty reduction has increased in recent years and will continue to do so [A good forecast] The concern with poverty reduction goes beyond the belief that economic growth will reduce poverty, to the view that targeted food subsidies as well as the medical and educational programs to which Williamson refers, can reduce the number of poor people . . and should be used for that purpose" (Williamson 1990 27)

6 Some people might wish to add nation-building to the noneconomic objectives to be pursued by development policy (as was common in the 1960s).

References

The word "processed" refers to informally reproduced works that may not be commonly available through library systems

Alam, M Shahid 1999 "Does Sovereignty Matter for Economic Growth?" In John Adams and Francesco Pigliaru, eds., *Economic Growth and Change.* Cheltenham, U.K. Edward Elgar

Balassa, Bela 1970 "Growth Strategies in Semi-Industrial Countries" *Quarterly Journal of Economics* 84(1):24–47

Brecher, Richard, and Carlos Diaz-Alejandro 1977 "Tariffs, Foreign Capital, and Immiserizing Growth" *Journal of International Economics* 7(4).317–22.

Burki, Javed, and Guillermo E Perry 1998. *Beyond the Washington Consensus Institutions Matter* Washington, D C.. World Bank

Geithner, Timothy F 1999. "The World Bank and the Frontier of Development Challenges" Remarks delivered to a conference on "Reinventing the World Bank Challenges and Opportunities for the 21st Century," Northwestern University, May 14 Processed

Hamada, Koichi. 1998. "IMF Special· Keeping Alive the Asian Monetary Fund" *Capital Trends* 3(September)

Hinkle, Lawrence E, and Peter J Montiel, eds 1999 *Exchange Rate Misalignment Concepts and Measurement for Developing Countries.* New York: Oxford University Press

Ito, Takatoshi. 1999. "The Role of IMF Advice." Paper presented to International Monetary Fund conference on "Key Issues in Reform of the International Monetary and Financial System," Washington, D C, May 29 Processed.

Kolodko, Gregorz. 1998. *Transition.* World Bank Development Economic Research Group newsletter, June.

Little, Ian, Tibor Scitovsky, and Maurice Scott. 1970. *The Structure of Protection in Developing Countries: A Comparative Study.* Baltimore, Md.: Johns Hopkins University Press.

McCloskey, Deirdre N. 1998. *The Rhetoric of Economics.* 2d ed Madison, Wis.: University of Wisconsin Press

Naím, Moisés. 1995. "Latin America: The Morning After." *Foreign Affairs* 74(July–August):45–6

———. 2000. "Washington Consensus or Washington Confusion?" *Foreign Policy* (spring):86–10

Rajan, Ramkishen S. 1999. "The Brazilian and Other Currency Crises of the 1990s." *Claremont Policy Brief.* Claremont College, Claremont, Calif. May.

Soros, George. 1998. *The Crisis of Global Capitalism.* New York: Public Affairs.

Stewart, Frances. 1997. "Williamson and the Washington Consensus Revisited." In Louis Emmeri ed., *Economic and Social Development into the XXI Century.* Washington, D.C.. Inter-America Development Bank.

Stiglitz, Joseph E. 1994. "The Role of the State in Financial Markets." In Michael Bruno and Bori Pleskovic, eds., *Proceedings of the World Bank Conference on Development Economics 1993* Washington, D.C.: World Bank.

———. 1998. "More Instruments and Broader Goals· Moving toward the Post–Washington Consensus." United Nations University/World Institute for Development Economics Researc Helsinki

———. 1999a "Whither Reform? Ten Years of the Transition." In Boris Plesovic and Joseph Stiglitz, eds., *World Bank Annual Conference on Development Economics* Washington, D.C.: Worl Bank.

——— 1999b "The World Bank at the Millennium " *Economic Journal* 109(459).F577–97.

United Nations 1998. *Annual Report of the United Nations University.* Tokyo.

Vines, David 1999 In an obituary for Susan Strange. Newsletter no 9 Economic and Social R search Council Programme on Global Economic Institutions, Swindon, U K Processed.

Waelbroeck, Jean 1998. "Half a Century of Development Economics: A Review Based on the *Han book of Development Economics " The World Bank Economic Review* 12(May):323–52.

Williamson, John. 1990 "What Washington Means by Policy Reform " In John Williamson, e *Latin American Adjustment· How Much Has Happened?* Washington, D C.: Institute for Intern tional Economics

———. 1996 "Are the Latin American Reforms Sustainable?" In Hermann Sautter and Rolf Schinl eds., *Stabilization and Reforms in Latin America· Where Do We Stand?* Frankfurt: Vervuert Verl

——— 1997 "The Washington Consensus Revisited." In Louis Emmerij, ed., *Economic and Soc Development into the XXI Century.* Washington, D C.: Inter-American Development Bank

——— 1999 "Implications of the East Asian Crisis for Debt Management " In A. Vasudevan, e *External Debt Management Issues, Lessons, and Preventive Measures.* Mumbai. Reserve Bank India

———. Forthcoming. *Exchange Rate Regimes for Emerging Economies· Reviving the Intermediate tion.* Washington, D.C.: Institute for International Economics

Williamson, John, and Molly Mahar 1998 "A Survey of Financial Liberalization." Princeton Ess in International Finance 211. Princeton University, Princeton, N.J.

World Bank. 1993. *The East Asian Miracle: Economic Growth and Public Policy.* New York. Oxf University Press.

——— 1996. *World Development Report: From Plan to Market* New York: Oxford University Pr

The Washington Consensus a Decade Later: Ideology and the Art and Science of Policy Advice

T. N. Srinivasan

In 1990 Williamson coined the term "Washington Consensus" to describe "the lowest common denominator of policy advice being addressed by the Washington institutions to Latin American countries as of 1989." He now protests in an article in this volume that the phrase has "become a synonym for 'neoliberalism' or what George Soros (1998) has called 'market fundamentalism'. . . ." This development should have caused no surprise, given the visceral hatred in many parts of the world for free markets. Because they view the U.S. government and its "lackeys"—the World Bank, the International Monetary Fund, and to a lesser extent, the Inter-American Development Bank—as the chief advocates of free markets, these opponents of free markets would have attacked anything called the "Washington Consensus." Had he instead called it the "Williamson synthesis," perhaps the package would have been a less inviting target for attack, although one cannot be sure: given the author's association with the hated Washington institutions, he could not have credibly absolved them from responsibility for the synthesis.

More than a decade has passed since Williamson proffered the consensus package. It is tempting (following the advice of the late Senator George Aiken of Vermont, who suggested that the United States simply declare victory and withdraw from the then-raging Vietnam War) to announce that the policy prescriptions contained in the consensus package have been successful and are no longer contentious! Even though there is considerable evidence of the success of many elements of the package where appropriately implemented, unfortunately the top echelon of at least one of the institutions celebrated in the consensus—the World Bank—seems to be misinformed and confused about the contents of the policy package. Others doubt the appropriateness of some of the recommended policies and argue that other elements of the package never commanded a consensus. It is useful, therefore, to reexamine

The World Bank Research Observer, vol. 15, no 2 (August 2000), pp 265–70

the package for its contemporary relevance, taking into account the experience of t 1990s.

At the outset, let me say that I regard the policy package less as a consensus a more as a reflection of the author's synthesis of lessons from four decades of dev opment experience. Indeed, Williamson's phrase "common denominator of pol advice" can mean only one thing: that the advice is based on policy implicatic that emerged not only from a number of studies of certain Latin American cot tries but also, more generally, from experience in many developing countries. this sense, trade liberalization as a sensible policy was supported by studies pu lished in the 1970s by the Organisation for Economic Co-operation and Develc ment (Little, Scitovsky, and Scott 1970), the National Bureau of Economic I search (Bhagwati 1978; Krueger 1978), and the World Bank (Balassa 1970). later World Bank study, Papageorgiou, Michaely, and Choksi 1991, lent simi support.) But only the utterly naive would interpret this advice as unqualified the sense of being appropriate regardless of the specific circumstances of an econo or the time horizon involved. Yet the studies covered many countries with varyi socio-political-economic institutions and many different time periods, and alm all of them still found possible benefits from trade liberalization—a fact that te fies to the robustness of the policy advice. Be that as it may, it is common kno edge among economists that the response to any policy change, such as trade lib alization, that operates through price incentives depends both on nonprice fact and on the time horizon. For example, if domestic supply constraints (other tl price received) are severe in the short and medium run, removing all price dist tions would have only a limited favorable response. And, to the extent that revenues are largely derived from trade taxes, the government may be constrai as to how far it can liberalize trade by reducing trade taxes without compromi fiscal discipline. This constraint may not be binding, however, if the actual app tariffs exceed revenue-maximizing levels, as is often claimed. In any case, one enumerate many nuances and caveats to each of the recommended policies in package. But only an ideologue or the utterly ignorant would conclude that cause caveats apply, any attempt to change the status quo through the implen tation of the recommended policies is undesirable.

Similarly, the advice to liberalize interest rates was a common conclusion several studies of financial repression in developing countries. Williamson now cedes that it could be costly to liberalize interest rates before other elements of fi cial liberalization, such as prudential supervision of banks by capable and kn edgeable central bank authorities, are in place. Although this point is obviously the scope of its applicability is arguable. For example, in many developing coun most banks are publicly owned, and whether they would or could gamble fc demption if deposit interest rates were to be liberalized is open to question. A

rate, as in the case of trade liberalization, the case for interest rate liberalization is likely to remain intact for many developing countries.

The advice on fiscal discipline was also based on the experience of many developing countries, particularly Latin American countries, that had previously undergone episodes of hyperinflation and stop-and-go sequences of stagnation and growth. India's experience in the 1980s, when it abandoned fiscal discipline to run deficits financed by costly borrowing at home and abroad, is instructive. The spurt in growth following the reckless fiscal expansionism proved unsustainable and ended inevitably in a macroeconomic-cum-balance of payments crisis in 1991. The fact that India did not experience Latin-style inflation before or after the abandonment of fiscal discipline is beside the point: because most of India's poor workers, particularly those in agriculture and informal service sectors, are not protected against inflation, even moderate inflation by Latin standards has serious consequences for the welfare of the poor in India.

The advice to redirect public expenditure toward health care, primary education, and infrastructure has long been part of conventional wisdom. A point that is not part of the conventional wisdom (and one that leading advocates of redirection such as Amartya Sen and his acolytes do not emphasize) is that failure to liberalize trade, to privatize inefficient public enterprises of dubious social value (such as airlines, hotels, and steel mills), and to reform the tax system eats away public resources that could otherwise be directed to the social sectors. Thus trade liberalization, privatization, and tax reform, which constitute three of the ten policies in the consensus package, are important not only in and of themselves but also because they make more resources available for social sectors.

The package called for "a competitive exchange rate," an unfortunate choice of words. An exchange rate is the price of one currency in terms of another currency or a basket of currencies. The word "competitive" used in the context of an exchange rate evokes painful memories of competitive devaluations by many countries in the interwar era. Under the classic Bretton Woods system of fixed exchange rates, one could interpret an "uncompetitive" exchange rate to mean one that is overvalued relative to its long-run equilibrium value. The operational significance of the interpretation is vastly diminished, however, by the fact that the long-run equilibrium rate is hard to define, that it was not defined in the Articles of Agreement of the International Monetary Fund, and that it is not simple to compute from available data in any case. Perhaps by a competitive exchange rate Williamson meant an undervalued exchange rate and that in fact Japan and the East Asian economies following Japan maintained competitive exchange rates in this sense.

Edwards and Savastano (1999) recently surveyed the empirical studies on exchange rate regimes in developing countries. They identified two camps: one group ascribes a key role to the exchange rate as a nominal anchor; the other stresses the perils of

relying on an asset price and therefore sees the exchange rate as a nominal anchor in a world of integrated global capital markets. In their view

> . . . the differences between the two broad camps identified lie in their differing views regarding three key features of exchange rate policy in a context of high capital mobility: (1) the scope for (and effectiveness of) sterilized and unsterilized intervention as a means for attaining (and preserving) a degree of nominal exchange rate stability; (2) the costs that "excessive" fluctuations of the nominal exchange rate may impose on the economy's performance, and (3) the *time dimension* of their analysis—i.e., the horizon over which monetary policy, the exchange rate, capital flows and the rest of the economy are assumed to interplay. All of these are empirical issues for which little, if anything, is known for the case of developing countries—not even for the relatively advanced ones. (Edwards and Savastano 1999:22)

Since the Asian financial crisis, a consensus seems to be emerging that the onl two viable exchange rate regimes are either a system of rigidly fixed exchange ra implemented through a currency board arrangement or its opposite, a regime freely floating exchange rates. Williamson, however, now prefers an intermedia regime between the two with limited flexibility. He does not explain why and ho the limited flexibility of the regime could be credibly signaled to distinguish it, the one hand, from the old-fashioned and now discredited crawling peg and, on t other hand, from a regime of transition to a free float. The advice to keep the e change rate "competitive" has no operational content even if a competitive exchan rate could be defined in conceptual terms. Of course, allowing the rate to float fre obviates this problem.

The recommendation to liberalize flows of foreign direct investment grew in p out of the need to have capital inflows that did not create debt and in part from t desire for other benefits such as technology transfers, which were believed to associated with such investments. Such investment flows are not likely to be as vo tile as short-term capital flows because a decision to invest in another country probably based on the long-term fundamentals of the recipient economy. Even late as 1989, private capital flows, including foreign direct investment, had not celerated as much as they did in the 1990s. The financial crises starting with Mex in December 1994 and ending with the most recent one in Brazil in 1998 h exposed several weaknesses in the domestic financial sectors of developing count and in the international financial architecture. The crises have, if anything, re forced the advice to liberalize foreign direct investment flows.

The advice to deregulate, in the sense of abolishing barriers to entry and, equ important, to exit, continues to be pertinent because it is based on age-old and pro virtues of competition. Indeed, benefits from removing price distortions will be limi

f firms not strong enough to compete in an undistorted market are not allowed to xit and if more efficient firms are denied entry.

The obverse side of deregulation is the need to ensure competition to firms that re privatized. For internationally traded goods and services (except possibly in wide-ody passenger jet aircraft!), the world market is far larger than the minimum effi-ient scale of production, so that opening to trade is adequate to generate competi-ion. In nontraded sectors where considerations of scale economies or network xternalities preclude the possibility of more than a handful of firms operating on an fficient scale, a regulatory authority needs to be established and an appropriate set f regulations promulgated so that the few private firms operate in a socially desir-ble manner. Interestingly, technological developments have vastly eroded scale econo-nies in electricity generation, telecommunications, and other nontraded goods and ervices that were once deemed natural monopolies.

In 1990, when Williamson published his consensus, regulation and privatization n developing countries were scarce, and there was none in Russia and the Eastern European countries, which were still part of the disintegrating Soviet Union. In the lecade since then, privatization has taken off, particularly in the transition econo-nies. It is fair to say, however, that the state of knowledge about appropriate mecha-iisms for privatization and regulatory frameworks (particularly for private financial ntermediaries) is still in a state of flux. Nonetheless, the advice to privatize remains elevant, with a cautionary note about the need to ensure sufficient competition for rivatized enterprises.

My brief evaluation of the Washington Consensus 10 years after its promulgation trongly suggests that its policy advice remains largely intact. In general, sound policy advice, while undoubtedly based on received theory and empirical evidence, neces-sarily has to involve judgment. Economic theorizing involves simplification and ab-straction of complex reality; econometric analysis of empirical evidence often im-poses restrictions on theory such as, for example, specific functional forms for utility and production functions, the nature of heterogeneity among firms and consumers, and distributional assumptions about stochastic terms. Simple and abstract theory (which in its "second-best" version could amount to saying that almost anything is possible!) and its highly restrictive econometric specification cannot deliver policy conclusions that can be directly applied to the situation of any given economy at a particular time. To advise on policy requires sound judgment on the part of the advisor—judgment that goes beyond findings for theoretical and econometric mod-els. Such judgment will incorporate knowledge about history, particularly economic history, and about the specific socio-political-cultural features and institutions of the country involved. For example, cross-country regressions, even if they are not mind-less, cannot deliver policy conclusions about the desirability of trade liberalization or capital inflows or about the effect of openness on growth. At best, such regressions are an efficient means for discovering patterns in the data from which capable re-

searchers can draw inferences after bringing to bear their knowledge about the econo mies involved while firmly eschewing attribution of causality. To say that judgmen based on the specifics of the case is needed is not to argue either that there can be n robust policy conclusions of wide applicability or that only discretion, rather tha rules, should govern policy choice. What it means is that formal analysis has to b supplemented by informally allowing for factors that by necessity have been exclude from the formal analysis. There is art as well as science in policy advising! An hones policy advisor will clearly indicate where his judgment enters and where theory an econometrics stop. It is not hard to isolate the theoretical and econometric bases an the astute judgment of Williamson in the Washington Consensus.

Note

T. N. Srinivasan is Samuel C Park Jr Professor of Economics at Yale University.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems

Balassa, Bela 1970 "Growth Strategies in Semi-Industrial Countries" *Quarterly Journal of Economics* 84(1):24–47.

Bhagwati, Jagdish 1978 *Foreign Trade Regimes and Economic Development Anatomy and Consequences of Exchange Control Regimes* Cambridge, Mass · Ballinger

Edwards, Sebastian, and Miguel A Savastano 1999 "Exchange Rates in Emerging Economies. What Do We Know? What Do We Need to Know?" Revised version of a paper presented at a Stanford University conference on "Economic Policy Reform What We Know and What We Need to Know," September 17–19, 1998. Processed.

Krueger, Anne. 1978. *Foreign Trade Regimes and Economic Development Liberalization Attempts and Consequences* Cambridge, Mass · Ballinger

Little Ian, Tibor Scitovsky, and Maurice Scott 1970 *The Structure of Protection in Developing Countries A Comparative Study.* Baltimore, Md. Johns Hopkins University Press.

Papageorgiou, Demetris, Michael Michaely, and Armeane M Choksi, eds 1991 *Liberalizing Foreign Trade* Oxford, U K · Basil Blackwell.

Soros, George. 1998. *The Crisis of Global Capitalism.* New York: Public Affairs.